The international journal of science / 25 May 2023

nature



BENDING THE RULES

Foldable silicon wafers for flexible solar cells

Toxic labs
Why is science experiencing a mental-health crisis?

Detection revolution
How to harness the power of quantum sensors

Wasting effort
Molecular mechanism linked to muscle loss in chronic disease

VISIT…

LANZAROTE Caliente.com

Nature.2023.05.27

[Sat, 27 May 2023]

- This Week
- News in Focus
- Books & Arts
- Opinion
- Work
- Research
- Collections

This Week

- **Encourage whistle-blowing: how universities can help to resolve research's mental-health crisis** [23 May 2023]
  • Low pay, job insecurity, bullying and harassment all contribute to academic researchers reporting above-average levels of anxiety and depression. Institutions can improve working environments by looking at best practice elsewhere.
- **Boost African research in exchange for debt relief** [24 May 2023]
  • Research will satisfy Africa's demand for energy — but only if the continent is allowed to invest more in its science and innovation.
- **Against climate hypocrisy: why the IPCC needs its own net-zero target** [23 May 2023]
  • A robust strategy to slash the IPCC's carbon emissions would be a testbed for international climate policy — and serve as an example of effective action.

- EDITORIAL
- 23 May 2023

Encourage whistle-blowing: how universities can help to resolve research’s mental-health crisis

Low pay, job insecurity, bullying and harassment all contribute to academic researchers reporting above-average levels of anxiety and depression. Institutions can improve working environments by looking at best practice elsewhere.



Universities should rethink their reluctance to encourage whistle-blowing.Credit: Alamy

Researchers working in academia are more likely to experience anxiety and depression than are members of the population at large, as we report in a Feature investigating the mental-health crisis in science. The COVID-19 pandemic has taken its toll on researchers, as it has on many in wider society, but it is clear that a major factor common in academia is a toxic work environment.

A proliferation of short-term contracts, low salaries (particularly for early-career researchers), competitive working environments and pressure to publish are all contributors — but so are bullying, discrimination and harassment. Study after study has reported on the devastating effects that these behaviours can have, especially on under-represented groups such as women, people of colour, low-income students and members of sexual and gender minorities.



A mental-health crisis is gripping science — toxic research culture is to blame

Regrettably, the finding that bullying and harassment are widespread in academia is not new, as reporting by *Nature* and other journals show. But few academic leaders seem to be doing much to solve the problem. That's not to say they aren't listening — many, perhaps even most, are. A considerable number are trying to implement policies to improve campus well-being. But these efforts are not yielding positive results.

To do better, university administrations and governing bodies need to look beyond the campus for solutions — in particular, to industry. Ultimately, internal structures need to change to reflect how science is performed today. And modern systems of employee redress for when things go wrong — such as those that encourage whistle-blowing — need to be implemented.

At the root of the problem is the mismatch between the idea of a university as a refuge for lone geniuses and the collaborative nature of contemporary science. As Athene Donald, a physicist at the University of Cambridge, UK, reminds us in her book *Not Just for the Boys: Why We Need More Women In Science*, published earlier this month, much science today is a team effort. It frequently involves large numbers of researchers collaborating within and across laboratories, and often across countries and cultures. What matters is that teams work together, complementing each other's strengths and compensating for one another's weaknesses, always with an eye on completing the task at hand. Good science matters more than hierarchies. Donald also reminds us that studies show that discovery and innovation benefit when teams include people from historically under-represented groups, but that these contributions are all too frequently ignored or undermined (B. Hofstra *et al. Proc. Natl Acad. Sci. USA* **117**, 9284–9291; 2020).

Yet the research system still tends to put power in the hands of just one person, or a few people. When grants are awarded, the responsibility for them is generally given to principal investigators, not shared more widely among members of a research team. That often means that large numbers of early-career researchers — 50 or more in some fields — are under the supervision of a single principal investigator. In the same way, departments and faculties have a single person at the top. And in general, this 'officer class' comprises people from less-diverse backgrounds than those they are supervising.



Stress and uncertainty drag down graduate students' satisfaction

It doesn't have to be this way. When power, responsibility and autonomy are shared, people work more collegially, and step up when required. But if power is vested in one person, or only a few people, it can be misused to harass, undermine and bully others. All too often, this is what happens.

This is why universities are structurally ill-equipped to provide redress against poor behaviour. Although universities do have both grievance and whistle-blowing procedures (which protect people from being penalized or dismissed if they disclose wrongdoing), they often discourage anonymization for those drawing attention to inappropriate behaviours or actions. Some argue that it makes complaints more difficult to investigate. This is often true, but anonymous whistle-blowing provides a fairer way to obtain justice in an environment in which the person perpetrating harm is often also the person ultimately responsible for an individual's career prospects.

Universities need to take a good look at this matter and study how internal organizational structures could be reformed to better reflect how research is done today. A good first step would be to look to industry, because we know that researchers working in industry report higher job satisfaction than do their academic peers. At the same time, academic governing bodies should revisit mechanisms for whistle-blowing, including studying how this is done

in businesses, in the public sector and in non-profit organizations, and evaluating the strengths and weaknesses of the approaches used in each.

Universities in many countries have extensive connections to industry, through industry–academia collaborations, through knowledge-exchange offices and through the direct involvement of business people, who serve as professors and as members of governing boards. These relationships could all be leveraged to study ways to bring about positive change in academia.

The need to achieve positive change cannot be overstated. Doing nothing is not an option. If the next generation of researchers is as dissatisfied as our Feature suggests, then no less than the future of research and scholarship is at stake.

Nature **617**, 651 (2023)

doi: https://doi.org/10.1038/d41586-023-01703-9

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01703-9

- EDITORIAL
- 24 May 2023

Boost African research in exchange for debt relief

Research will satisfy Africa’s demand for energy — but only if the continent is allowed to invest more in its science and innovation.

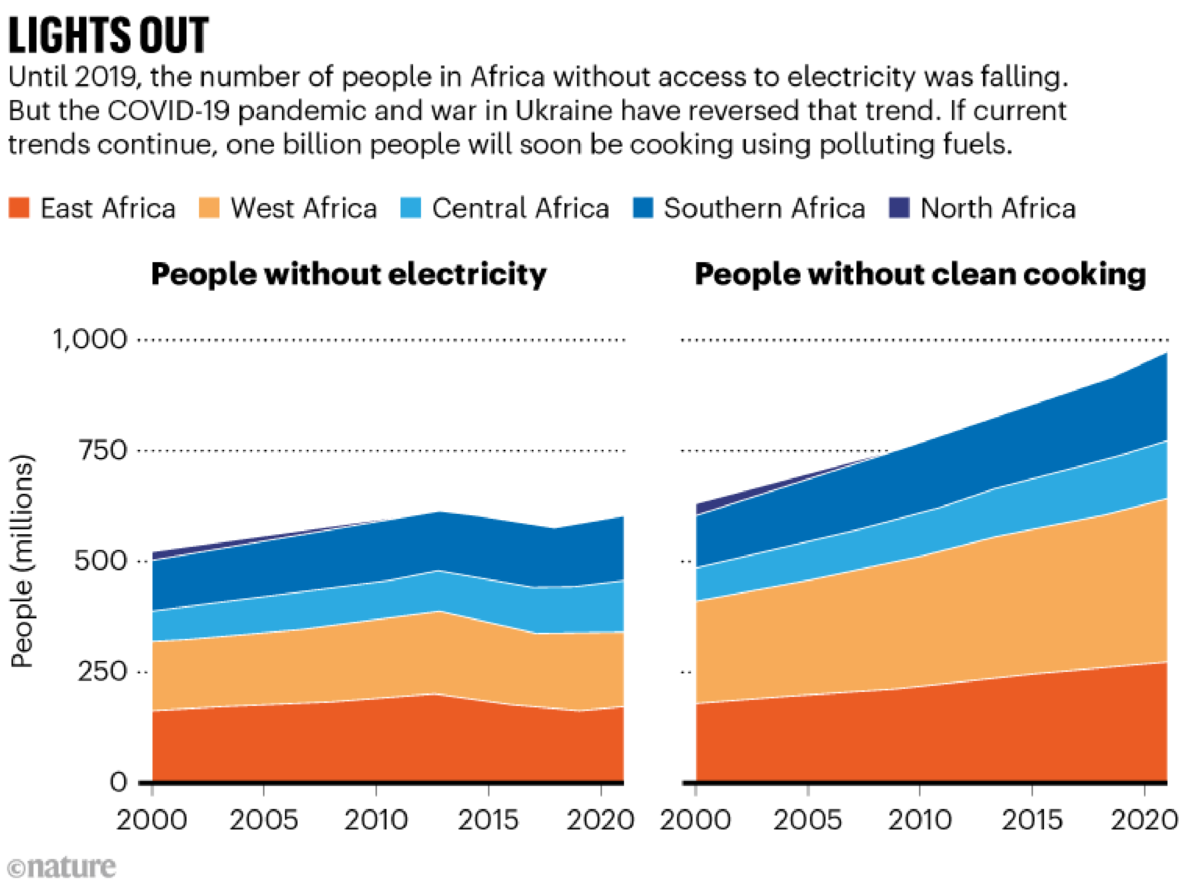


Solar-powered LED lights are helping young people in Africa to continue their studies after sunset, when mains-powered electricity isn’t available.Credit: Waldo Swiegers/Bloomberg/Getty

There’s one data point in the World Bank’s latest *Africa’s Pulse* report, published last month, that is impossible to ignore: 43% of the continent’s people still do not have electricity (see go.nature.com/3oploj). That number is staggering not only because it’s so high, but also because it represents a

rise since the COVID-19 pandemic began (see ‘Lights out’). We are living in a time when tens of billions of dollars are being spent on developing artificial intelligence, searching for personalized medicine and exploring space. But back on Earth, 600 million Africans are being denied the ability to turn on a light switch. And around 900 million people in sub-Saharan Africa are still cooking by burning charcoal and firewood (G. Bensch *et al.* *One Earth* **4**, 879–890; 2021).

The lingering effects of pandemic-induced lockdowns and disrupted supply chains are clearly contributing to the lack of these most fundamental services, the report says. But there’s arguably a bigger and more systemic reason: highly indebted countries are being stopped from investing their resources in research and development. And that, in turn, makes it harder for them to attract private investment.



Source: International Energy Agency

A report published last month by the African Development Bank shows that companies involved in mining Africa's cobalt, lithium and nickel for use in electric-vehicle batteries will make around US$11 billion in 2025 (see go.nature.com/3Iy80ob). However, those that manufacture and assemble these batteries — most, if not all, of whom are outside Africa — will earn more than $1 trillion. The report therefore recommends that African countries establish centres of research excellence to help them participate in this higher-value manufacturing. This is an excellent idea, and it should be supported — but it simply isn't an option for governments that are struggling to keep up with day-to-day spending.

By the end of last year, 22 African countries were in or at risk of debt distress. Their leading creditors, such as the International Monetary Fund (IMF) and governments, notably China, insist that debt repayments be prioritized, so they will need to spend most of what they earn through taxes on paying off debts. But the dysfunctional nature of today's geopolitics means that there's little coordination between these creditors. The IMF typically requires indebted nations to increase taxes on goods and services and decrease public spending. China's demands are rarely made public. The IMF argues that its loan conditions safeguard spending on education, health and social protection. But a report last month from the international aid group Oxfam says that countries find it impossible to both pay off debts and protect essential public spending (see go.nature.com/43alco).



The United States can help the IMF to rethink how it lends

All of this leaves highly indebted countries with little to spend on investments, including support for their research programmes. And that means that African countries will find it hard, if not impossible, to build long-term research and development capacity in areas such as renewable energy.

Over the past several months, Africa's leaders have held high-level meetings, organized by the United Nations, the African Union and individual governments, to discuss the need to boost investments in science and technology. The latest gathering, Africa Week, is taking place this week at the University of Pretoria in South Africa, in partnership with the country's government and National Research Foundation. It will focus discussions on open science and equitable international research partnerships. The ambition of more equitable partnerships must also include conversations about meaningful reductions in the continent's debts.

In the past, Africa's research policymakers have made the case for 'debt-for-science swaps' — in which creditors agree to waive some debt for countries that spend more on research and development. This is not a new principle. The idea of reducing debt in exchange for a longer term investment is already being applied by the IMF and other creditors to help tackle environmental issues. Countries have been offered debt forgiveness in exchange for action on climate change or conservation goals, for example. Africa's creditors should consider reviving this principle for research, too.

It's a scandal that 600 million people in Africa lack access to something as fundamental as electricity — and that this number is not declining. Without debt-for-science swaps and other creative solutions, too many of Africa's countries are at risk of being trapped in a doom loop of debt and austerity that hurts the poorest and most vulnerable populations.

Nature **617**, 652 (2023)

doi: https://doi.org/10.1038/d41586-023-01704-8

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01704-8

| Section menu | Main menu |

- WORLD VIEW
- 23 May 2023

Against climate hypocrisy: why the IPCC needs its own net-zero target



A robust strategy to slash the IPCC’s carbon emissions would be a testbed for international climate policy — and serve as an example of effective action.

- Benjamin M. Sanderson

In April, I flew to Bangkok for the final meeting of the United Nations Intergovernmental Panel on Climate Change (IPCC) 6th Assessment Cycle (AR6). My return flight to Stockholm emitted more carbon than any one person in the lowest-emitting 50% of the global population contributes in an entire year.

The latest AR6 synthesis report (see go.nature.com/4hfkzpc) concludes that rapid, deep and immediate emissions reductions are needed to maintain a

liveable world. But this powerful messaging is diluted by the IPCC’s actions.

The IPCC should aggressively limit its own emissions instead of requiring in-person sessions and the attendant long-haul flights. Although meetings contribute only a tiny fraction of total global emissions, improving accountability would have an outsized impact on the IPCC’s effectiveness, and would be a case study for robust, internationally coordinated mitigation.

The COVID-19 pandemic has demonstrated that many IPCC meetings can be conducted fully remotely, and most climate researchers are in favour of virtual elements in conferences. This is in line with trends in the wider climate-science community: the World Climate Research Programme’s Coupled Model Intercomparison Project, which helps to coordinate climate models, has proposed that it should lead by example and transition to net-zero emissions as fast as possible. But similar statements from the IPCC are notably absent.



Net-zero emissions targets are vague: three ways to fix

This irony does not go unnoticed across the political spectrum. Conservatives point to individual researchers’ emissions to justify lack of urgency or to validate their view that institutional decarbonization is

impossible; progressives see a lack of effective action and a growing philosophical barrier between science and activism. Researchers' individual actions resonate with the public: in one US study, support for carbon taxation increased by up to 60% when respondents were told that the researchers had made efforts to reduce their carbon footprints (S. Z. Attari *et al. Clim. Change* **154**, 529–545; 2019).

The in-person model also acts as a demographic and social filter. The need to spend extended time away from home can lead to gender bias: a survey of academic faculty members indicated that mothers are more than twice as likely as fathers to experience parenting stress as a result of attending conferences (see go.nature.com/3o9ds). In-person meetings also create barriers owing to difficulties obtaining visas. And requiring travel risks under-representing the generation that will experience the most severe climate impacts, because non-fliers, a growing fraction of scientists, are more likely to be younger.

There are wider concerns about vaguely defined net-zero policies, which allow policymakers to broadcast their intentions to meet climate targets sometime in the future, while global emissions remain broadly constant today. Emissions-reduction plans are highly dependent on the future development and large-scale deployment of technologies to remove carbon dioxide from the air. The technological and physical plausibility of achieving net zero, as well as its definition, are subjects of extensive scientific debate.



[An analysis of ways to decarbonize conference travel after COVID-19]

By taking aggressive action to reduce emissions during the next assessment cycle (which ends in 2030), the IPCC could be a test case for a better approach. It could first outline an objective and unbiased process for quantifying current emissions, then define and assess the risks of different pathways for complete decarbonization. The emissions associated with the IPCC process are not trivial, but they are manageable. The challenges mirror those of the wider mitigation problem: assessment of direct and indirect emissions, the need for fairness in the face of robust targets, the reliability of removal methods and the need for international agreement on how to verify them.

Emissions reductions should not come at the cost of the IPCC's ability to assess and communicate cutting-edge science. On the one hand, some in-person meetings might be needed to facilitate communications on sensitive or thorny topics, but on the other, off-the-record backroom chats can entrench positions of influence.

Such decisions need to be made in a cost–benefit framework, informed by an accurately quantified carbon footprint. For instance, if the IPCC is to use CO_2 removal to address its remaining unavoidable emissions and reach net zero, it must be transparent about the inherent uncertainties in the

permanence of the carbon-storage technologies used. Such efforts would increase global understanding and industrial competence in CO_2 removal while acknowledging that the technology can be effective on a global scale only if paired with radical emissions cuts.

Even failure to achieve targets would be useful, by highlighting real-world limitations in net-zero policy assumptions, which could then inform wider societal strategy.

As the July elections for the next assessment cycle approach, IPCC members should remember that nobody is better placed to demonstrate how to eliminate emissions. The IPCC's actions matter, not just its words. Although it might make climate scientists uncomfortable, what the IPCC does about its own carbon emissions might be as crucial to its effectiveness as advancing cogent and robust science.

Nature **617**, 653 (2023)

doi: https://doi.org/10.1038/d41586-023-01707-5

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01707-5

| Section menu | Main menu |

News in Focus

- **The world's top chemical-weapons detectives just opened a brand-new lab** [17 May 2023]
  • The state-of-the-art centre will help to enforce a near-universal ban on certain chemicals and train analysts from around the world.
- **Carbon capture key to Biden's new power-plant rule: is the tech ready?** [11 May 2023]
  • Regulation aims to reduce facilities' emissions of climate-altering gases — a goal that could require carbon burial.
- **How one man's rare Alzheimer's mutation delayed the onset of disease** [15 May 2023]
  • Genetic resilience found in a person predisposed to early-onset dementia could potentially lead to new treatments.
- **Why is COVID life-threatening for some people? Genetics study offers clues** [17 May 2023]
  • Immune genes could play a part in the risk of needing intensive care when infected with SARS-CoV-2.
- **'The disease will be neglected': scientists react to the WHO ending mpox emergency** [12 May 2023]
  • Researchers worry that certain regions, including Africa, will suffer if global attention moves away from the outbreak.
- **Hammerhead sharks are first fish found to 'hold their breath'** [11 May 2023]
  • It pays to be an warm hunter in the cold ocean depths, so the animals shut down oxygen intake to conserve heat.
- **China's mysterious spaceplane returns to Earth — what we know** [18 May 2023]
  • Specialists speculate that it might be similar to a US spaceplane, and it could have research or military uses.
- **A mental-health crisis is gripping science — toxic research culture is to blame** [23 May 2023]

• With researchers reporting high rates of anxiety and depression, calls are growing to fundamentally change science before it’s too late.

| Next section | Main menu | Previous section |

- NEWS
- 17 May 2023

The world’s top chemical-weapons detectives just opened a brand-new lab

The state-of-the-art centre will help to enforce a near-universal ban on certain chemicals and train analysts from around the world.

- Davide Castelvecchi



OPCW inspectors are trained to use hazmat suits with a self-contained breathing apparatus.Credit: OPCW/Flickr (CC BY-ND 2.0)

The Hague, the Netherlands

The international body that banned chemical weapons is due to celebrate its first major milestone sometime this year — the completed destruction of the world's declared stockpiles of banned substances. But at the organization's brand-new facility in the Netherlands, scientists from around the world will continue its work to prevent, spot and respond to chemical warfare.

On 12 May, the Organisation for the Prohibition of Chemical Weapons (OPCW) officially inaugurated its new Centre for Chemistry and Technology near The Hague, where the international body will bring together its existing laboratories and add new monitoring and training programmes.



Attacks in UK and Syria highlight growing need for chemical-forensics expertise

Although nearly all countries — except Egypt, Israel, North Korea and South Sudan — have officially renounced the use of chemical weapons under an international treaty called the Chemical Weapons Convention (CWC), such chemicals continue to be used illegally. The OPCW has played a central part in the response to crises: its investigators found that nerve gas and chlorine were used in the Syrian civil war during the past decade, for example, and in 2018 an OPCW team investigated the attempted

assassination of a Russian dissident in the United Kingdom that involved a nerve agent called Novichok.

Historic step forward

The centre “will allow the OPCW to enhance the [lab] verification regime through maintaining and developing our knowledge, skills and expertise related to chemical weapons”, says OPCW director-general Fernando Arias.

“It was an historic step forward for the OPCW today,” said Paul Walker, a vice-chair of the board at the Arms Control Association, a non-governmental organization in Washington DC, after attending the 12 May launch. Although the OPCW’s monitoring work had been going on for more than a decade, the Syrian civil war catapulted it onto the global stage. “That began the whole process of realizing that we needed [the OPCW to have] much more expansive and expensive labs,” says Walker, who attended the inauguration on behalf of an umbrella organization called the CWC Coalition.



The new Centre for Chemistry and Technology is an upgrade compared with previous OPCW lab facilities.Credit: OPCW/Flickr (CC BY-ND 2.0)

“I think it will be a big deal for the OPCW,” says Alastair Hay, an environmental toxicologist at the University of Leeds, UK, who has been an OPCW adviser and conducts outreach training for the organization. “The previous lab facilities OPCW had were really quite limited,” he says. The lab will also enhance the organization’s ability to train researchers, he adds. The centre was built outside the OPCW’s normal budget, with €34 million (US$37 million) of voluntary contributions from 56 countries.

Stockpile disposal

Since 1997, the OPCW has had a mandate under the CWC to monitor countries’ compliance with the ban, which required signatories to stop producing chemical weapons, to declare their existing stockpiles and commit to destroying them. The organization has so far certified the destruction of more than 70,000 tonnes of chemical weapons, and it received the Nobel Peace Prize for its work in 2013. Of the eight countries that disclosed the possession of such weapons, seven have completed disposal. The eighth — the United States — expects to finish destroying its stockpile (mainly artillery containing mustard gas and nerve agents) in the next few months and is thought to be the last country to do so.

“We estimate that a large number of old and abandoned chemical weapons still remain undiscovered around the globe,” says OPCW spokesperson Elisabeth Waechter. For example, large numbers of rusting Japanese chemical weapons left over from the Second World War still await safe disposal in China. “While it is the responsibility of countries to destroy them safely, OPCW technical experts can provide advice — and verify the safe destruction.”

The CWC also compels state parties to report the production of chemicals that have a dual purpose — compounds that have legitimate uses but can also be used either as weapons or as ingredients to make them.



The OPCW analyses samples from chemical plants worldwide to search for evidence of the production of banned chemicals.Credit: OPCW/Flickr (CC BY-ND 2.0)

At the new centre, a warehouse is stocked with crates containing analytical-chemistry equipment ready to be dispatched around the world. As well as investigating alleged chemical-weapons attacks, the OPCW sends inspectors on routine missions to chemical facilities in member countries, often at short notice, to search for evidence of the production of banned chemicals. “Inspectors go to different sites and essentially audit what those particular sites have declared,” says Hay.

Samples brought back to The Hague are also sent to at least two international labs, to ensure that the results are cross-checked before they are reported to member countries. “It’s really important to be able to have access to really good facilities, where the results are incontrovertible,” Hay notes.



[Novichok nerve agents banned by chemical-weapons treaty](#)

These facilities are part of a network of 25 ‘OPCW-designated’ labs. The remit of the Centre for Chemistry and Technology includes putting labs through a periodic certification process to enable participation in the network. The centre also trains chemists from labs that want to match its standards and join the list. In an effort to provide knowledge to as many nations as possible, the OPCW is working on certifying labs in African countries for the first time: its current trainees include scientists from Algeria, Morocco, Kenya, Nigeria and South Africa.

Increasing the labs’ geographical diversity can “build global capacity for the benefit of all regions and the international community as a whole”, Waechter says.

Nature **617**, 657-658 (2023)

doi: https://doi.org/10.1038/d41586-023-01622-9

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01622-9

- NEWS
- 11 May 2023

Carbon capture key to Biden’s new power-plant rule: is the tech ready?

Regulation aims to reduce facilities’ emissions of climate-altering gases — a goal that could require carbon burial.

- Jeff Tollefson



Some suspect that the new power-plant regulation proposed by the Biden administration is really designed to shut down coal-fired plants.Credit: Jim Urquhart/Reuters

The administration of US President Joe Biden has proposed a landmark regulation that aims to curb emissions from the power sector in the United States over the next two decades. If it survives the legal challenges that are sure to come from power-plant operators and others, the rule could hasten an ongoing shift towards renewable energy in the country. It could also, perhaps, help to jump-start the use of a long-stalled — and controversial — technology that enables utilities to capture and bury carbon emissions from power plants fired by fossil fuels.



US aims for electric-car revolution — will it work?

The much-anticipated regulation from the US Environmental Protection Agency (EPA), released on 11 May, lays out a range of proposed requirements for power plants fuelled by natural gas and coal, which are currently responsible for more than 1.5 billion tonnes of carbon dioxide (CO_2) emissions annually in the United States — roughly one-quarter of the country's climate pollution. Large coal-fired power plants that plan to operate beyond 2040 would need to use carbon-capture and sequestration technology (CCS) to reduce their emissions by 90%, for instance; major gas-fired plants could deploy either CCS or transition to clean hydrogen (produced with minimal carbon emissions) as a fuel source. The EPA will accept comments on the proposal for 60 days.

Overall, the agency says that the rule could curb US emissions of CO_2 by more than one billion tonnes by 2042 — equivalent to the annual carbon emissions of Japan — while racking up as much as US$85 billion in climate and health benefits. “The public health and environmental benefits of this proposed rule will be tremendous,” said EPA administrator Michael Regan at a press conference on 10 May, and it will have minimal impacts on electricity prices.

The power-plant regulation arrives on the heels of an ambitious rule proposed by the Biden administration to reduce climate pollution from cars. Many scientists and environmentalists hoped that it would put teeth in the administration’s commitment to decarbonize the US power sector by 2035, but the regulation would allow fossil-fuel power plants to continue emitting well beyond that date.



How a dangerous stew of air pollution is choking the United States

“This is a good step and a good framework, long overdue but very welcome,” says David Doniger, who heads the climate and clean air programme at the Natural Resources Defense Council, an advocacy organization based in New York City. Nonetheless, Doniger says that his group will push the EPA to implement tougher measures in the final regulation.

Whether the rule holds up to court challenges will hinge in part on the argument that CCS is ready for prime time. Many experts say it is — particularly with new tax incentives intended to drive down the cost of the technology. "The biggest barrier to widespread use of CCS in the power industry today is economics," says Howard Herzog, an engineer who studies CCS at the Massachusetts Institute of Technology in Cambridge.

Inside the fence line

The EPA proposal follows a key US Supreme Court ruling last June, which held that the agency had overstepped its authority under former president Barack Obama by crafting broad regulations that would have pushed the whole power sector towards cleaner energy sources, such as wind and solar. By contrast, the new EPA rule focuses on emissions reductions that can be achieved "inside the fence line" at individual power plants, with the best available technologies. These include CCS.



Environmental Protection Agency administrator Michael Regan has said that the new rule would benefit public health.Credit: Bonnie

Cash/UPI/Shutterstock

This is a more conventional regulatory approach to limiting pollution than what the Obama administration attempted. It also fits in with the Supreme Court decision last year, which acknowledged the EPA's authority to regulate emissions at the scale of individual power plants, says Julie McNamara, an energy analyst with the Union of Concerned Scientists, an advocacy group based in Cambridge, Massachusetts.

One of the core questions will be whether CCS is an economically viable technology — or whether the new EPA rule has actually been designed to push electric utilities away from fossil fuels and towards cleaner energy sources. Regan directly addressed this concern at the press conference. He acknowledged that some coal-fired power plants would close as a result of the rule, but added that "this is really a decision that will be made state by state, company by company".

West Virginia attorney-general Patrick Morrisey has already issued a statement expressing dissatisfaction. The new rule, he said "just seems designed to scare more coal-fired power plants into retirement". West Virginia has a large coal industry and last year helped to bring the case against the EPA to the Supreme Court. "We expect that we would once again prevail in court against this out-of-control agency," he said.

McNamara says that the current EPA proposal is just one piece of a larger puzzle that includes other regulations on hazardous pollutants, as well as massive public-health incentives for low-carbon energy. The question facing regulators, utilities and communities will be what makes the most sense: investing in CCS technologies to prolong the life of a fossil-fuel industry with enormous impacts on climate and public health, or investing in clean energy? "We feel very strongly clean energy will be the superior choice," McNamara says.

Technological viability

Although CCS technology has been demonstrated at only a handful of power plants around the world, interest is rising. This is particularly the case in

countries, such as China, that depend heavily on coal, the dirtiest of fossil fuels, according to Jarad Daniels, who leads the Global CCS Institute, an advocacy group based in Melbourne, Australia, with offices in Washington DC. CCS is also moving forwards in sectors such as cement, steel and chemicals.



US Supreme Court hobbles the EPA's authority over climate emissions

EPA officials say that the technology is ready to implement at a reasonable cost — or even a profit — depending on the plant. Many scientists and environmentalists agree. "If you ask, 'why hasn't this technology been more widely used in the past', it's because there are no regulations, and it's cheaper to keep dumping emissions into the atmosphere for free," Doniger says.

In a spending bill enacted last year, the US Congress boosted federal incentives for CCS in the power and other industrial sectors to US$85 per tonne of CO_2, which experts say mostly offsets the cost of capturing the gas and pumping it underground. However, economic modelling by multiple groups suggests that this is still not enough to convince most electric utilities to embrace the technology, given falling costs for renewable-energy sources such as solar.

“The economics only pencil in for CCS in very specific settings,” says Ben King, an energy analyst with the Rhodium Group, a consultancy based in New York City. But clean-energy sources could face their own challenges in scaling up quickly, owing to potential bottlenecks with licensing, securing sites for facilities and delivering power over long-distance electric transmission lines. An analysis by the Rhodium Group suggests that regulatory requirements might yet enable CCS to gain a foothold in the power sector.

How things play out will vary by plant. Electric utilities will probably choose to close down older coal-fired power plants, for instance, whereas some new gas-fired plants that run frequently could be candidates for a CCS retrofit, says John Thompson, an energy analyst with the Clean Air Task Force, an advocacy group based in Boston, Massachusetts. The scales could be tipped by a combination of economic incentives that lower the cost, and regulations that require action, he adds.

“What you get is synergies between the rule that says, ‘Thou shalt do this’, and the tax credit that says ‘This isn’t very expensive’,” Thompson says. “The potential is there for deep reductions in emissions.”

Nature **617**, 658-660 (2023)

doi: https://doi.org/10.1038/d41586-023-01595-9

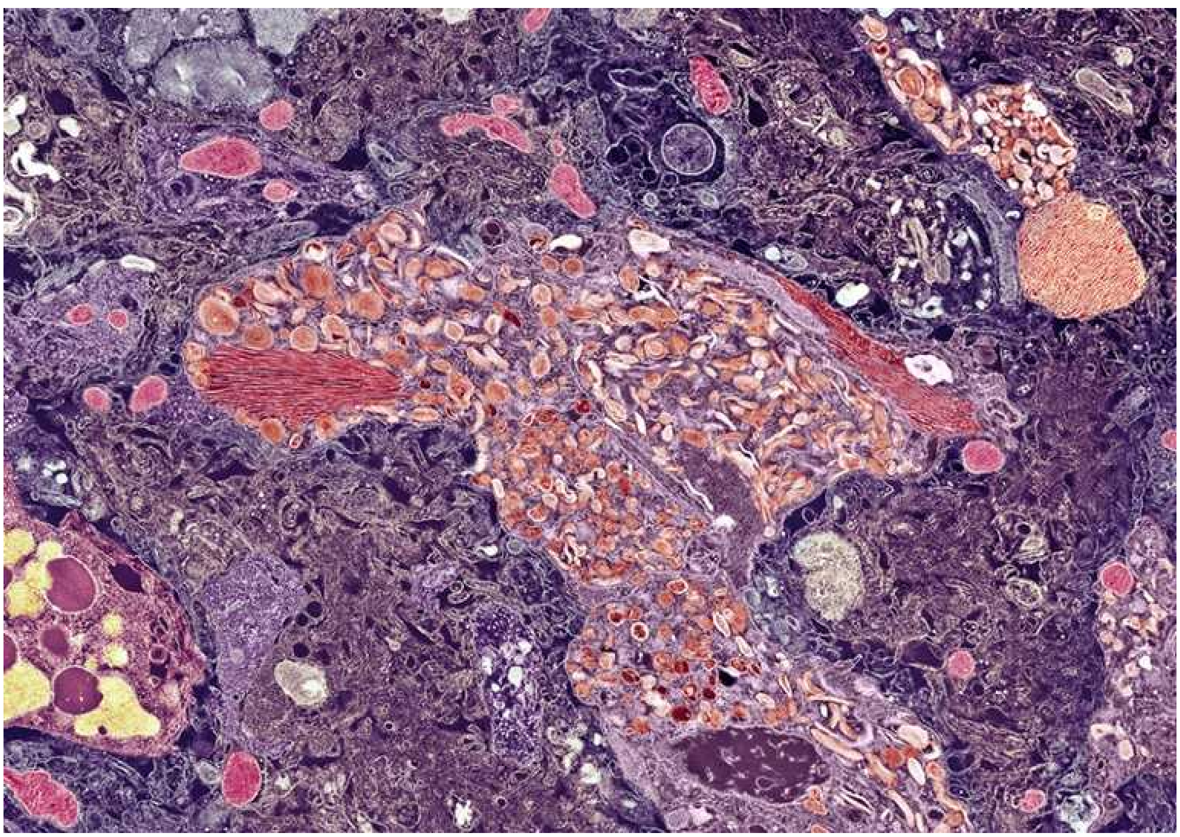
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01595-9

- NEWS
- 15 May 2023

How one man’s rare Alzheimer’s mutation delayed the onset of disease

Genetic resilience found in a person predisposed to early-onset dementia could potentially lead to new treatments.

- Sara Reardon



Electron-microscope image showing an amyloid plaque and tau fibres (dark pink lines) in the brain of a person with Alzheimer's disease. Amyloid plaques are insoluble aggregates of amyloid-β protein.Credit: Thomas Deerink/NCIR/SPL

Researchers have identified a man with a rare genetic mutation that protected him from developing dementia at an early age. The finding, published on 15 May in *Nature Medicine*[1], could help researchers to better understand the causes of Alzheimer's disease and potentially lead to new treatments.

For nearly 40 years, neurologist Francisco Lopera at the University of Antioquia in Medellín, Colombia, has been following an extended family whose members develop Alzheimer's in their forties or earlier. Many of the approximately 6,000 family members carry a genetic variant called the paisa mutation that inevitably leads to early-onset dementia. But now, Lopera and his collaborators have identified a family member with a second genetic mutation — one that protected him from dementia until age 67.

"Reading that paper made the hair on my arms stand up," says neuroscientist Catherine Kaczorowski at the University of Michigan in Ann Arbor. "It's just such an important new avenue to pursue new therapies for Alzheimer's disease."

Mutated protein

Lopera and his colleagues analysed the genomes and medical histories of 1,200 Colombians with the paisa mutation, which causes dementia around ages 45—50. They identified the man with the second mutation when he was 67 and had only mild cognitive impairment.

When the researchers scanned his brain, they found high levels of the sticky protein complexes known as amyloid plaques, which are thought to kill neurons and cause dementia, as well as a protein called tau that accumulates as the disease progresses. The brain looked like that of a person with severe dementia, says study co-author Joseph Arboleda, an ophthalmologist at Harvard Medical School in Boston, Massachusetts. But a small brain area

called the entorhinal cortex, which coordinates skills such as memory and navigation, had low levels of tau.

The researchers found that the man had a mutation in a gene coding for a protein called reelin, which is associated with brain disorders including schizophrenia and autism. Little is known about reelin's role in Alzheimer's, so the researchers genetically engineered mice with the same mutation. In mice, the mutated form of reelin caused the tau protein to be chemically modified, limiting its ability to cluster around neurons.

The study challenges the theory that Alzheimer's disease is driven primarily by amyloid plaques, which are the targets of several drugs recently approved by the US Food and Drug Administration. The drugs effectively remove amyloid from the brain, but lead to only a moderate improvement in rates of cognitive decline.

The fact that the man stayed mentally healthy for so long despite the many amyloid plaques in his brain suggests that Alzheimer's is more complicated, says Yadong Huang, a neurologist at the Gladstone Institutes in San Francisco, California. He suggests that there could be multiple subtypes of Alzheimer's, only some of which are driven by amyloid. “We do need different pathways to really finally deal with this disease,” he says. The link to tau, he says, is especially promising because it suggests that tau has a role in mental decline. Several therapies targeting tau are currently in clinical trials.

Shared mechanisms

Lopera says that the reelin mutation is extremely rare in the general population, but that his team is now looking for this and other mutations in people with the paisa mutation. The man's sister, who had both the paisa and reelin mutations, began developing cognitive impairment at age 58 and severe dementia at 64 — later than average for someone with the paisa mutation. The authors say that she had experienced head injuries and had other disorders that could have contributed to her developing dementia earlier than her brother.

Arboleda notes that the mutated reelin protein binds to the same receptors as a protein called APOE, which is also associated with Alzheimer’s disease in people who do not have the paisa mutation. In 2019, the same group had identified a woman with the paisa mutation who developed dementia 30 years later than average, owing to a mutation in APOE[2]. Like the man in the latest study, the woman’s brain contained much higher levels of amyloid than would be expected in someone with so few Alzheimer’s symptoms.

“It’s really cool because it’s telling us there’s some shared mechanisms,” Kaczorowski says. Reelin and APOE compete to bind to the receptor, and the two findings suggest that either a stronger reelin protein or a weaker APOE protein can protect the brain against the disease. Arboleda says this suggests that therapies targeting reelin or APOE might be even more effective in sporadic Alzheimer’s cases, which tend to be less aggressive and progress more slowly than the early-onset type that the Colombian family experiences.

As with many people with Alzheimer’s, the man’s hippocampus — a brain region controlling learning and memory — was smaller than average at the time of his death, suggesting that it was degenerating. But because his cognitive abilities remained relatively intact, Kaczorowski says, neurons in other parts of the brain might have repurposed themselves to make up for the damage. Knowing whether that happens, she adds, could help to inform future therapeutic strategies.

“The vast majority of research focuses on why some people have Alzheimer’s; very few are on conditions where a factor can go against this disease,” says Huang. He says that further research is needed to pin down the mechanism through which reelin and APOE affect tau, and whether targeting these proteins could help people with Alzheimer’s who do not have the paisa mutation. “This is one of those few cases that really opens the door for anti-Alzheimer’s research.”

Nature **617**, 660-661 (2023)

doi: https://doi.org/10.1038/d41586-023-01610-z

References

1. Lopera, F. *et al. Nature Med*. https://doi.org/10.1038/s41591-023-02318-3 (2023).

2. Arboleda-Velasquez, J. F. *et al. Nature Med.* **25**, 1680–1683 (2019).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01610-z

- NEWS
- 17 May 2023

Why is COVID life-threatening for some people? Genetics study offers clues

Immune genes could play a part in the risk of needing intensive care when infected with SARS-CoV-2.

- Heidi Ledford



Many people with COVID-19 required hospitalization during the pandemic; a study now reveals genetic links between severe illness and activation of

the immune system.Credit: Ariana Drehsler/AFP via Getty

An analysis of DNA from more than 24,000 people who had COVID-19 and required treatment in intensive care has yielded more than a dozen new genetic links to the risk of developing extreme illness from the disease.

The study, which was published on 17 May in *Nature*[1] and has more than 2,000 authors, highlights the role of the immune system in fuelling the later stages of particularly severe COVID-19. The results could one day contribute to the development of therapies for COVID-19 — and potentially other diseases that cause acute respiratory distress or sepsis.

"These are likely to be processes that are active in other conditions," says Kenneth Baillie, an intensive-care specialist at the University of Edinburgh, UK, and lead author on the study. "Everything that we've done in COVID will, I think, be relevant to other groups of patients that we haven't identified yet."

Team effort

Much of the data were collected from people who were infected with SARS-CoV-2 during the first waves of the pandemic in the United Kingdom. At that time, many intensive-care staff members were terrified of spending time on the ward, says Baillie. A coronavirus outbreak in 2003 caused by the related virus SARS-CoV had a mortality rate of about 10% and had hit health-care workers particularly hard. "We were nervous," says Baillie. "Certainly, in the first wave of the pandemic, and for some people in subsequent waves, it seemed like a considerable personal risk."

Despite their fear and the frenzy of providing care at the onset of the pandemic, scores of health-care workers took on the added burden of enrolling participants in studies that investigate how genetics might have contributed to disease severity. Many of the participants were on ventilators, and enrolling them meant long discussions with relatives who were going through some of the most difficult times of their lives, says Baillie. The result was an effort called GenOMICC, which researchers hope will help to

improve treatment options for COVID-19 and other conditions in future pandemics.



[How your first brush with COVID warps your immunity]

Baillie and his colleagues analysed data from more than 24,000 people and combined this information with data sets from around the world. They found 49 DNA sequences that are associated with becoming critically ill from COVID-19. Sixteen of these had not been reported previously.

Among these sequences are some that could affect the activity of genes and proteins involved in the immune system. Raging immune cells have long been implicated in causing some of the tissue damage seen in late-stage, severe cases of COVID-19. Baillie and his colleagues found genetic links to inflammatory responses and the activation of immune cells — processes that can damage the lungs and reduce their capacity to send oxygen to the body's tissues.

"It definitely expands our understanding of the genetic determinants of severe COVID," says Brent Richards, a human geneticist at McGill University in Montreal, Canada. Richards is an investigator on another project called the COVID-19 Host Genetics Initiative, a global effort in which scientists from more than 54 countries share data.

The diversity of that effort will provide a crucial counterpart to the GenOMICC study, which is predominantly drawn from participants of European descent, says Richards.

Reliable results

In addition to the new findings, it is important that the UK study has replicated many of the findings by other groups, says Alexander Hoischen, a geneticist at Radboud University Medical Center in Nijmegen, the Netherlands. Replication of results across different data sets is crucial when trying to tease apart the possible contribution of common genetic variants to disease.

Even so, it will take a great deal of work to show that these variants contribute directly to critical illness — a necessary step to turn these initial findings into new therapies. "That is still a long trajectory," says Hoischen.

But putting together these large consortia of investigators and taking the first steps are considerable triumphs, he adds. In the early days of the pandemic, it was unclear how participation in such large endeavours would benefit the careers of individual scientists. "Many of them acted as purely human beings," he says, instead of worrying about their careers. "It illustrates the degree of solidarity that we really witnessed in COVID times."

Nature **617**, 661-662 (2023)

doi: https://doi.org/10.1038/d41586-023-01655-0

References

1. Pairo-Castineira, E. *et al. Nature* https://doi.org/10.1038/s41586-023-06034-3 (2023).

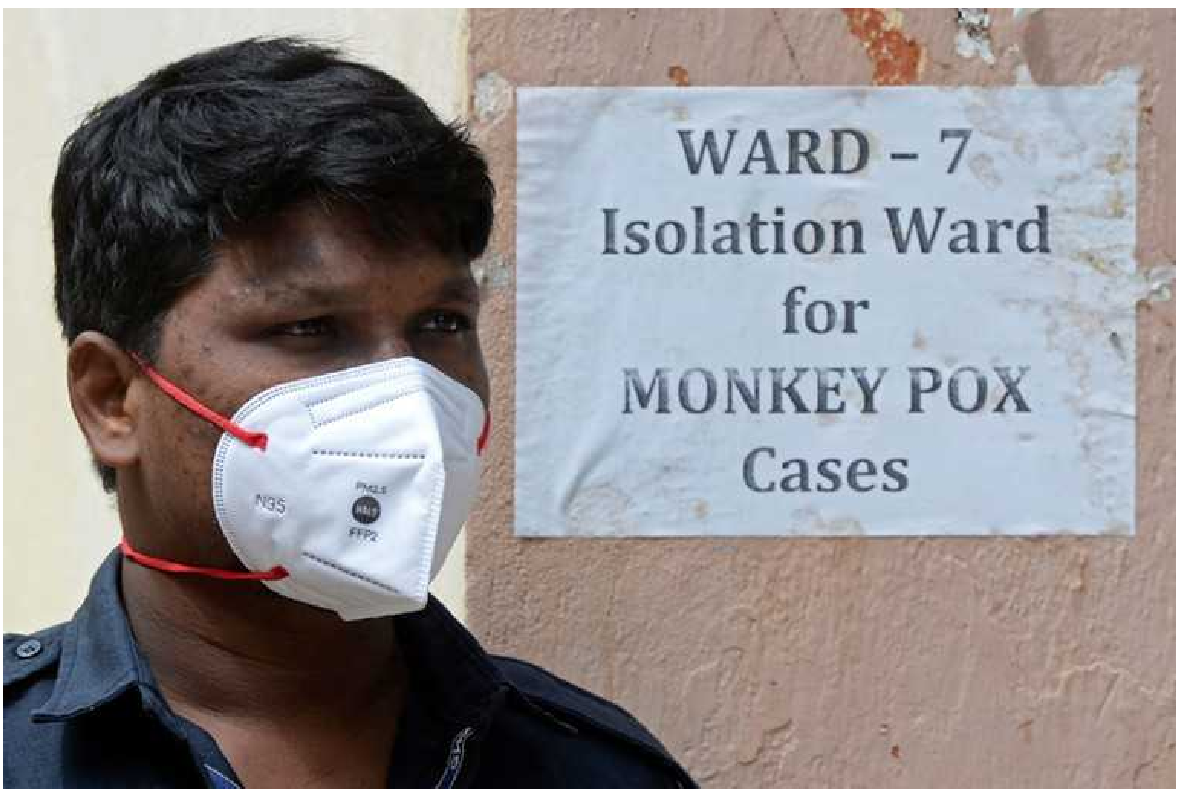
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01655-0

- NEWS
- 12 May 2023

‘The disease will be neglected’: scientists react to the WHO ending mpox emergency

Researchers worry that certain regions, including Africa, will suffer if global attention moves away from the outbreak.

- Layal Liverpool



Cases of mpox — the disease previously known as monkeypox — are still rising in some parts of the world.Credit: Noah Seelam/AFP/Getty

The World Health Organization (WHO) announced on 11 May that mpox is no longer a public-health emergency of international concern (PHEIC) — but researchers fear that the decision will draw crucial resources away from curbing the disease, which is still prevalent in regions including Africa.

"My biggest worry is that investments and resources for tackling mpox will dwindle, and efforts triggered by the 2022 outbreak will not be sustained," says Boghuma Titanji, a physician-scientist in infectious diseases at Emory University in Atlanta, Georgia. Tedros Adhanom Ghebreyesus, director-general of the WHO, based in Geneva, Switzerland, made the decision after a recommendation by the organization's mpox emergency committee on 10 May.

"We have many problems we are facing in Africa," says Dimie Ogoina, an infectious-disease physician at Niger Delta University in Amassoma, Nigeria. Ogoina is a member of the WHO emergency committee, but he didn't agree with its consensus recommendation to end the PHEIC. "The disease will be neglected just the way it has been neglected for a long time before the 2022 outbreak," he says.

Ogoina and Titanji say that the declaration of mpox as a PHEIC in July 2022 helped to focus the world's attention on the disease, which had previously been neglected despite being endemic to several African countries. (Until last year, the disease was called monkeypox, and the virus continues to be known by that name.)



Monkeypox in Africa: the science the world ignored

Millions of vaccines have been distributed to tackle the outbreak, but mainly in rich countries such as the United States and the United Kingdom. Vaccines and treatments have mostly been out of reach of African nations.

“Before the 2022 outbreaks and mpox becoming a disease of global importance, it was a disease that disproportionately affected poor populations in remote parts of West and Central Africa,” says Titanji. “I fear that we will see a return to that status quo at least until perhaps the next outbreak that impacts wealthy countries in the West.”

Disease of priority

The PHEIC designation helped to stimulate support from governments and public-health leaders in Nigeria and other affected countries on the continent, says Ogoina. “It also has a way of making African leaders, politicians, public-health authorities see that monkeypox is a disease of priority,” he says.

In his announcement, Tedros pointed to global progress in controlling the mpox outbreak in recent months, noting that there were almost 90% fewer cases reported in the past three months compared with the previous three.

However, the WHO’s latest situation report, published on 11 May, shows a rise in cases globally in the past fortnight.

Since the the WHO’s 27 April situation report, 264 new mpox cases and 10 related deaths have been reported to the organization. Most of the infections were in the Americas, particularly in Mexico and Brazil, and in the western Pacific region, including in Japan, China, South Korea, Singapore and Australia. In total, more than 87,000 mpox cases and 140 deaths have been recorded in 111 countries since the outbreak began in 2022.

Titanji says that the coming Northern Hemisphere summer might encourage infection spread. “We’re beginning to hear about reports of new clusters of cases, some of these cases being reported in people who were vaccinated last summer — and there’s still lots of things we don’t understand about this virus and how it’s moving within populations,” says Titanji. “These questions need to be addressed before we take a victory lap and say, we’re done with this.”

“Large swathes of the population do not have immunity against mpox,” she adds. “If the disease is introduced into the right networks by travel, by festivals and all that happens around the summer, then it’s not unrealistic that we may yet again see surges in this outbreak.”

Incomplete reporting

The latest WHO report notes that, in Africa, trends in cases remain unclear owing to reporting inconsistencies. “We’re not sure of the level of ongoing transmission because what we have in Africa is essentially an undercount,” says Ogoina. He also fears that the number of infections will rise in the coming months. “We’re going into a period where we may have social gatherings that may serve as superspreader events. There are still ongoing transmissions in some countries although there has been significant decline. My concern is that we may have a resurgence.”

Ogoina underscores that African nations still have far fewer resources for controlling infections than do rich countries. “PHEIC was an opportunity for us to galvanize support to see that we understand the disease better in Africa,

to see that we have vaccines and therapeutics in Africa, but currently we still don't understand the disease as we ought to," he says. "We still don't have vaccines. We still don't have therapeutics."

When asked by *Nature* at a press conference on 11 May if the decision to end the PHEIC classification would result in mpox being neglected, Michael Ryan, the WHO's executive director of health emergencies, said: "It is a neglected disease. And it's been totally neglected during this outbreak. In fact, WHO had to fund all of this international response purely on the basis of a contingency fund for emergencies — not one dollar was received from donors to support this response and support countries."

"We don't want this to be a neglected disease going forward," added Rosamund Lewis, the WHO's technical lead for mpox. "Efforts need to be made to continue surveillance, to continue capacity building in African countries specifically, to continue working with most affected communities." Those groups include men who have sex with men and people who don't have access to testing, vaccines or treatments, she said.

Nature **617**, 662-663 (2023)

doi: https://doi.org/10.1038/d41586-023-01581-1

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01581-1

- NEWS
- 11 May 2023

Hammerhead sharks are first fish found to ‘hold their breath’

It pays to be an warm hunter in the cold ocean depths, so the animals shut down oxygen intake to conserve heat.

- Bianca Nogrady



A juvenile scalloped hammerhead shark (*Sphyrna lewini*), with its mouth and gills open, near the ocean surface off Hawaii.Credit: Biosphoto/Alamy

Because it makes them better hunters, scalloped hammerhead sharks (*Sphyrna lewini*) have evolved a unique method to avoid losing body heat when they dive for prey in deep, cold waters: they close their gills.

Numerous fish and marine-mammal species are known to dive from the warm surface to deeper waters to hunt. However, ectothermic, or ‘cold-blooded’, animals face the challenge of how to conserve their body temperature to keep their metabolism active enough for hunting when the surrounding water can be just a few degrees above freezing.

“The most rapid point of heat loss for any fish, even a high-performance fish, is always at the gills,” says Mark Royer, a postdoctoral researcher in shark physiology and behaviour at the Hawaiʻi Institute of Marine Biology in Kaneohe, part of the University of Hawaiʻi. Because of the high volume of warm blood flowing through the gills, they are “essentially just giant radiators strapped to your head”, he says.

Some fish, such as the whale shark (*Rhincodon typus*), are able to conserve their body heat when diving through sheer size. Others, such as tuna, marlin and the family that includes great white (*Carcharodon carcharias*) and mako sharks (*Isurus oxyrinchus*), have evolved specialized heat-exchange systems at the gills that avoid too much body heat being lost.

The scalloped hammerhead has neither of these advantages or adaptations, yet has been tracked doing rapid, repeated dives to depths of around 800 metres.

To understand how sharks were coping with the temperature changes, Royer and his colleagues developed a device consisting of instruments that measured depth, water temperature, location and movement, as well as a probe embedded into muscles near the dorsal fin that recorded the shark’s core temperature. The device was designed to break off after several weeks, float to the surface and send out a signal to enable its recovery.

Three scalloped hammerhead sharks captured off the Hawaiian coast were tagged with the device.

In a paper published in *Science*[1], the team reported that the sharks would dive several times — six in an evening, for one shark — into deep water at temperatures of 5–11 °C, around 20 °C colder than at the surface, and remain there for 5–7 minutes at a time before surfacing.

Body temperature remained constant for most of the dive until the final stage of their ascent back to warmer waters, when it would decline rapidly.

Keeping warm

Royer suggests that the sharks are keeping their core temperature stable by simply not opening their gills or mouth during the dive; effectively ‘holding their breath’. “If you don’t have water going over your gills, then you won’t be dumping your body heat into the environment,” he says.

The drop in body temperature near the surface coincides with the sharks’ steep ascent flattening out slightly, which Royer suggests is the moment they start to allow water to flow over the gills. “They can slow themselves down, open their gills and start breathing again [because] the water that they’re in, it’s not as cold as it is at the bottom,” he says.

To shut down oxygen intake in this way suggests that the scalloped hammerhead must be able to deal with plummeting blood-oxygen levels during dives, says Mark Meekan, a fish ecologist at the University of Western Australia Oceans Institute in Perth, although the mechanism is yet to be discovered. “What they could be doing is slowing the heart muscle, slowing the pumping of blood around the body,” Meekan says. The shark’s tissues and blood could have evolved to hold more oxygen per unit of volume — akin to the adaptation seen in people who live at altitude — or might be able to deal with the molecular by-products of anaerobic respiration, which can be toxic at high levels.

Marine biologist Colin Simpfendorfer at James Cook University in Townsville, Australia, says the study shows how sharks are well adapted to the limits of their environment. “Diving to over 1,000 metres from tropical temperatures at the surface down to just a couple of degrees centigrade to

feed is a fairly extreme movement to do on a regular basis,” Simpfendorfer says.

So far, scalloped hammerheads are the first fish found to do this, but Simpfendorfer says other sharks and fish might have the same adaptation. “There is a big advantage in being able to hunt when you’re warm and everything else is cold.”

Nature **617**, 663 (2023)

doi: https://doi.org/10.1038/d41586-023-01569-x

References

1. Royer, M. *et al. Science* https://doi.org/10.1126/science.add4445 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01569-x

- NEWS EXPLAINER
- 18 May 2023

China’s mysterious spaceplane returns to Earth — what we know

Specialists speculate that it might be similar to a US spaceplane, and it could have research or military uses.

- Yvaine Ye



The Boeing X-37B, pictured here, is thought to have inspired China’s mysterious spaceplane.Credit: Boeing/Sipa/Shutterstock

China has landed a mysterious spacecraft that spent nine months in Earth's orbit, making the nation one of just a handful of entities that have successfully operated a reusable spacecraft. Although China has not revealed details about the spacecraft's design and operation, aerospace engineers have pieced together information from fragments, painting a picture of a craft with potential for research as well as military uses.

"The complete success of this experiment marks a significant breakthrough in our country's research on reusable spacecraft technology," Chinese state media agency Xinhua reported, suggesting that Chinese researchers have developed technology, such as heat shields and landing equipment, specifically for reusable craft.

China, the US companies Boeing and SpaceX and NASA are the only players to have launched and recovered orbiting reusable craft. Although the upfront costs are high, the costs of running reusable craft are a fraction of those for single-use craft, making reusable spacecraft attractive if a mission requires several flights. The nine-month mission is probably a second test of the reusable spacecraft. In September 2020, China returned a reusable experimental spacecraft that had spent two days in orbit.

Imitating Boeing?

The spacecraft is likely to be similar to a US spaceplane, the Boeing X-37B, according to Kevin Pollpeter, a senior research scientist at the Center for Naval Analyses, a US government-funded research and analysis organization in Arlington, Virginia. The purpose of X-37B is unknown, but he says the Chinese government was concerned about the craft's military capacities when it was revealed in 2010.

Jonathan McDowell, an astrophysicist at Center for Astrophysics Harvard & Smithsonian in Boston, Massachusetts, agrees that China's experimental spacecraft could be a nod to Boeing's X-37B spaceplane. McDowell says satellite sensing shows the spacecraft gliding and performing a runway landing at the Lop Nur military base in China's northwest Xinjiang region, suggesting that it is a plane. Some other reusable spacecraft, such as

SpaceX's Dragon capsule, are modules attached to launch rockets that rely on parachutes to land.

Like the X-37B, the Chinese vehicle seems to be fairly small. The Chinese media reported that the spacecraft was launched last August from the Jiuquan Satellite Launch Center in northwestern China on a Long March 2F rocket, which China used to launch its Shenzhou crewed space missions. Given that the maximum payload capacity of the launch vehicle is 8.4 tonnes, the spaceplane probably weighs between 5 and 8 tonnes, says McDowell — much smaller than NASA's retired space shuttles that performed crewed missions, including *Columbia* and *Challenger*.

Although the current model is too small to carry crew members, McDowell says he would not be surprised if the country introduces a larger spaceplane for crewed missions in the future.

Satellite capture?

China might be testing a variety of technologies that could further its space ambitions, says McDowell. Changing orbital trajectories or unfolding solar panels are just two examples of the kinds of mid-orbit manoeuvre that would be useful to perfect.

Releasing and capturing satellites while in orbit might also have been tested. According to a report by the Center for Strategic and International Studies, a think tank in Washington DC, the spaceplane released an object that was detected last October. The object seemed to disappear from orbit in January before reappearing in March, according to McDowell, who tracks the spacecraft using US Space Force data. One possible explanation, he says, is that the spaceplane captured the object and carried it as cargo before releasing it again. That would suggest the spaceplane has cargo capacity, meaning it could be used to carry instruments and satellites. The manoeuvre, which he says could have been achieved with a robotic arm, might have been practice for satellite maintenance and repair.

"[The Chinese] have been working a lot with robot arms in other contexts, like the Chinese space station and other satellite servicing experiments,"

McDowell says. For example, last year China removed a defunct satellite from orbit using another satellite.

The Xinhua story about the spacecraft's landing highlighted "the peaceful use of space". McDowell notes that, with plenty of spy satellites already in space, China doesn't necessarily need a spaceplane for that purpose. But China's space programme is run largely by the military, and consequently, its space missions tend to be secretive, says Pollpeter. "I think the inclination is to release less information rather than more information, because even if they are using it for something like satellite servicing and repair, it doesn't mean that it couldn't be used for more offensive or more military activities."

Nature **617**, 664 (2023)

doi: https://doi.org/10.1038/d41586-023-01639-0

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01639-0

- NEWS FEATURE
- 23 May 2023

A mental-health crisis is gripping science — toxic research culture is to blame

With researchers reporting high rates of anxiety and depression, calls are growing to fundamentally change science before it's too late.

- Shannon Hall



Illustration by Sébastien Thibault

There is a mental-health crisis in science — at all career stages and across the world. Graduate students are being harassed and discriminated against, paid meagre wages, bullied, overworked and sometimes sexually assaulted. It doesn't get much better for early-career researchers struggling to land long-term employment. And established senior researchers face immense pressure to win grants, publish in high-profile journals and maintain their reputations in highly competitive fields.

Scientists have raised concerns for years about the impacts of all these pressures on mental health. But a series of studies in the past few years are now providing hard data. And the findings show that the situation is dire.

Researchers are much more likely than the general population to experience depression and anxiety. And although the COVID-19 pandemic caused an increase in mental-health struggles, many argue that it only exacerbated problems that were already present. The recent studies, which have collectively surveyed tens of thousands of researchers worldwide, suggest that scientists' mental-health struggles are a direct result of a toxic research culture.



Science careers and mental health

That is particularly true for members of under-represented groups, including women, non-binary individuals, people of colour, those from sexual and gender minorities (LGBTQ+) and students on low incomes. But they also affect senior researchers and scientists in different countries.

“Well-being and how to set healthy boundaries in your life and in your work is a fundamental skill,” says Sharon Milgram, director of the Office of Intramural Training and Education at the US National Institutes of Health (NIH). “I feel like there was a blind spot in myself and many of us, in that it took this data to wake us up.”

With hard numbers in hand, some argue that science is at the beginning of a movement — one that will encourage systemic changes to improve the mental health of researchers over generations to come. Others argue that change is happening too slowly for young scientists who are already fleeing science — an effect that could have grim consequences for the future of research and society itself.

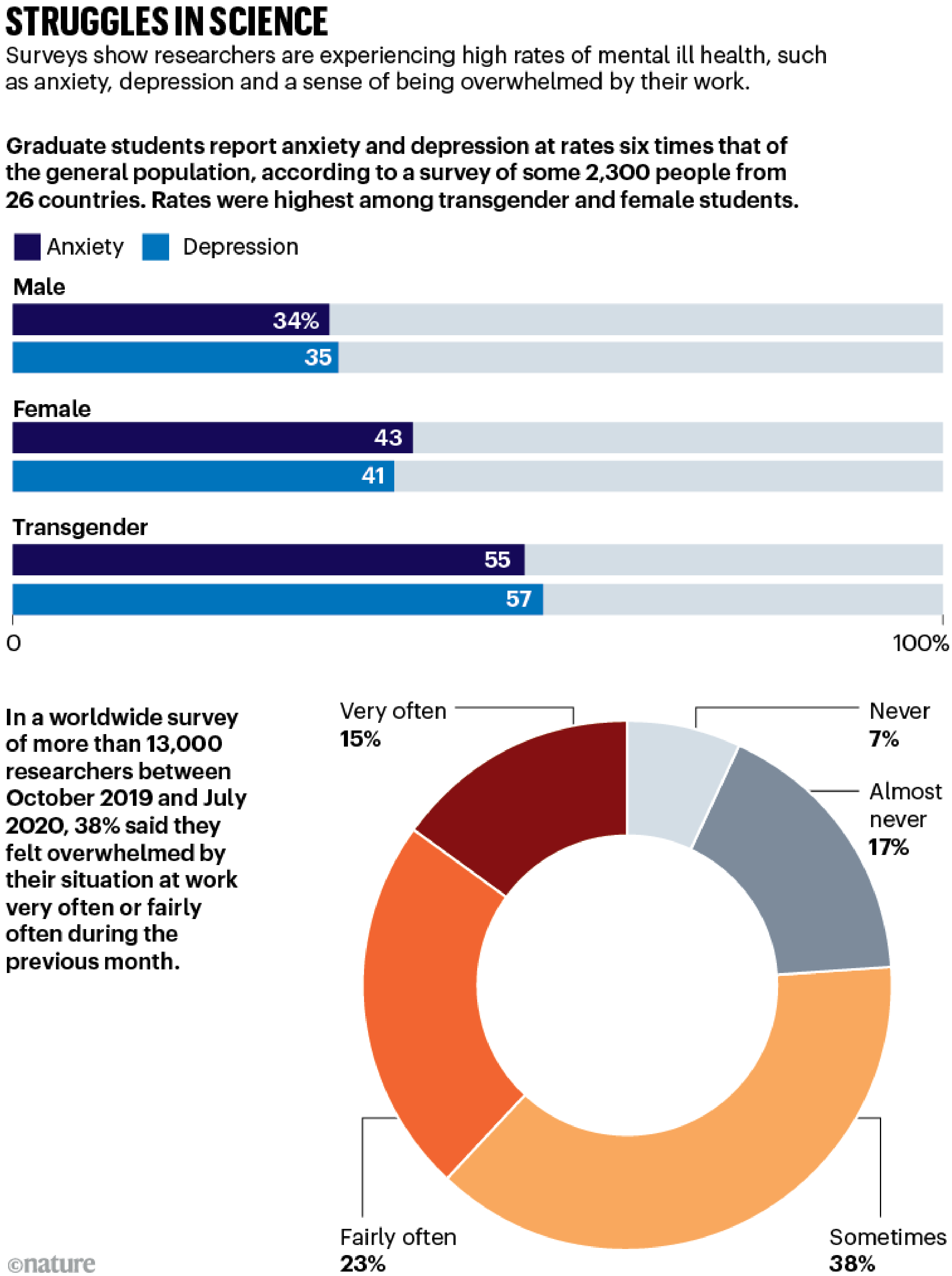
A global problem

In 2015, Teresa Evans, who directed graduate biomedical career development at the University of Texas Health Science Center at San Antonio, learnt that her students were struggling. But when they came to her for advice, she felt ill-equipped to help.

So Evans started her own research, only to uncover a dearth of literature on the topic. Not only were there few resources on how to help students, but it was unclear how extensive mental-health problems were — compelling Evans to circulate her own survey to quantify the matter. She received 2,279 responses, mostly from PhD candidates, at 234 institutions across 26 countries.

The results, published in March 2018, represented the largest survey of its kind at the time. It revealed a global problem: 41% of respondents reported moderate to severe anxiety and 39% had moderate to severe depression[1]. Those levels are six times greater than in the general population (see ‘Struggles in science’). The data also suggested possible sources of these

mental-health problems — anxiety and depression were often correlated with poor work–life balance and poor mentor relationships.

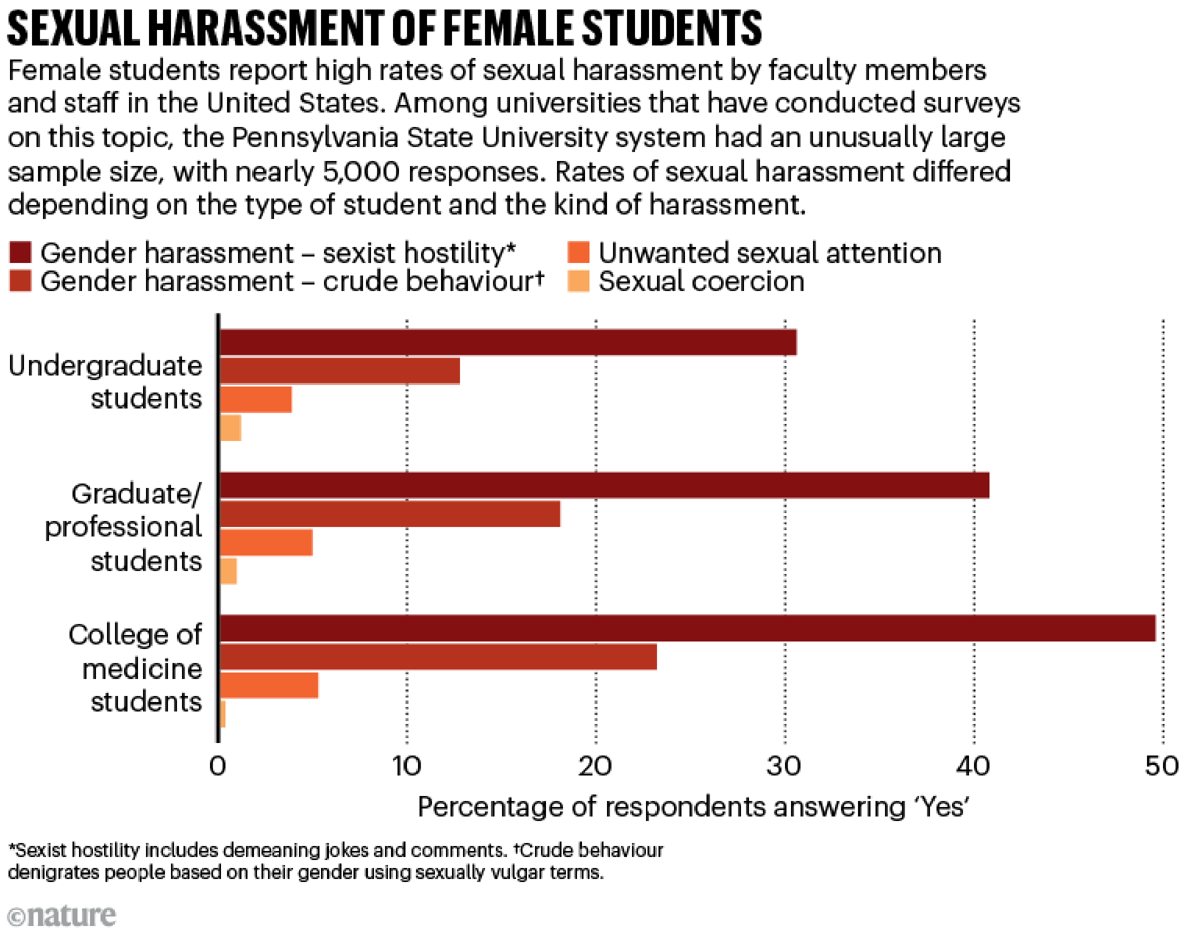


Sources: Left, Ref. 1; Right: Ref. 9

The rates varied significantly by gender: female, transgender and gender-nonconforming respondents were more likely to struggle with mental health than were their male counterparts. The prevalence of anxiety and depression was 55% and 57% for transgender and gender-nonconforming graduate students, 43% and 41% for women and 34% and 35% for men. That didn't surprise Evans, because women are more prone to anxiety and depression than are men[2].

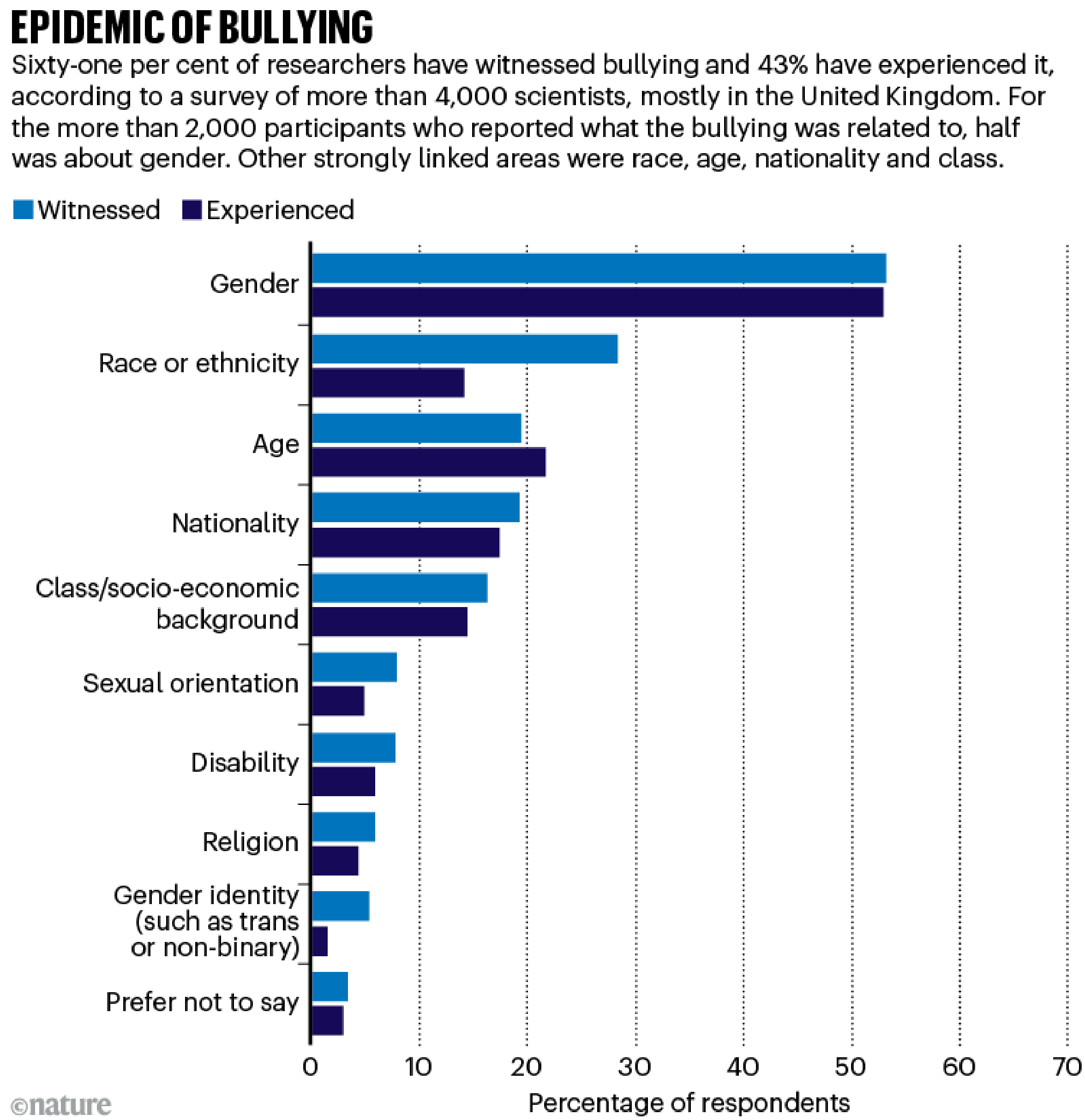
But there was another, bigger factor at work: sexual harassment in science that disproportionately affects women. In 2014, Kathryn Clancy, an anthropologist at the University of Illinois in Urbana-Champaign, documented high rates of sexual harassment in field science[3] . Then, in 2017, she and her colleagues surveyed 474 astronomers and planetary scientists, and found that 30% of women felt unsafe because of their gender (compared with 2% of men)[4].

These studies encouraged the US National Academies of Sciences, Engineering, and Medicine to appoint a special committee to examine the issue in academic settings. In June 2018, it released a report that revealed pervasive and damaging sexual harassment in science[5] (see 'Sexual harassment of female students').



Source: Ref. 5

It's not just a problem in the United States. In 2020, Wellcome, a major biomedical research funder in London, surveyed more than 4,200 scientists from 87 countries across genders, career stages and disciplines, and found that 43% of participants had experienced bullying or harassment and 61% had witnessed it[6] (see 'Epidemic of bullying'). Many felt that it was "culturally systemic", and 33% thought that leaders turned a blind eye to the behaviour.



Source: Ref. 6

In a 2021 *Nature* survey of more than 3,200 working scientists, nearly one-third said they had observed discrimination against or harassment of colleagues in their current job. The following year, another *Nature* survey of more than 3,200 PhD and master's students found that 35% of scientists who identify as members of minority racial or ethnic groups said they had experienced harassment or discrimination during their graduate studies.

It's no wonder there is a serious mental-health problem in science, say researchers. One study even found that the rates of burnout, depression and anxiety were comparable with those reported in 'high-risk' occupations such as health care[7].

The Wellcome study[6] found that 70% of respondents felt stressed on the average workday, and 34% had sought professional help for mental-health issues. Beyond harassment, many participants blamed funders and institutes that emphasize quantity over quality in terms of publishing and obtaining grants — all of which contribute to a poor work–life balance.

Another explanation suggested in the study[6] is that many scientists see their work as a vocation, not just a job. Although that means researchers are passionate, it also presents a unique challenge, argues Eric Pellegrini, an astronomer who left the field last year to become an independent data scientist. "They get you to torture yourself by making this work part of your identity — it's not a job, it's not even a career, it's a life choice," he says. "And you buy into that for years until you figure out what it is. Take away a lot of the parts, make it more generic, and it's just an abusive relationship."

Pandemic pressures

That was the picture before 2020. Then the COVID-19 pandemic hit, and with it, an onslaught of further challenges.

In a 2020 survey of 5,247 graduate students in science, technology, engineering and mathematics from 9 institutions across the United States, 38% reported symptoms consistent with anxiety and 35% had depression[8]. These proportions represented large jumps from what the same team found in 2019. The number of students with depression doubled, and the prevalence of anxiety rose by 50%[8].

"Our findings indicated that it was really a disaster," says Igor Chirikov, a senior researcher at the Center for Studies in Higher Education at the University of California, Berkeley, who led the study.



Postdoctoral researchers and other academics demand higher salaries and better benefits during a strike against the University of California last December.Credit: Sarah Reingewirtz/MediaNews Group/Los Angeles Daily News/Getty

Chirikov's team found that challenges in mental health were often associated with financial stress, which are acute for early-stage researchers. In the United States, for example, PhD students in the biological sciences earn salaries that are well below the cost of living. Postdoctoral researchers earn an average of $47,500 a year — a little more than half of the average annual salary for university graduates. Moreover, postdocs face constant upheaval, because they typically have to move to a new role every few years. Some researchers spend a decade or longer jumping from one short-term contract to the next.

Then, despite the years that scientists devote to training, many struggle to find a long-term job at a university, and this can drive young researchers out

of science entirely. The 2020 Wellcome survey[66] found that nearly half of the respondents who had left research reported that difficulty in finding a job was one reason. The two other common reasons included a negative impact on mental health and a desire for a better work–life balance.

Even if scientists land permanent positions, the competition never ends. In 2020, a survey designed by Cactus Communications, a science communication and technology company headquartered in Mumbai, India, analysed the opinions of 13,000 researchers in more than 160 countries. It found, for example, that 65% of respondents were under tremendous pressure to publish papers, secure grants and complete projects to maintain their reputation in the research community[9]. “That response is coming largely from more senior researchers, because they need to continue to be seen as bright as they once were,” says Abhishek Goel, co-founder and chief executive of Cactus Communications.

Beyond the pressure to perform, the study identified several other factors contributing to poor mental health among researchers, including long working hours and a culture in which stress and anxiety are normalized. Another factor mentioned by a large proportion of scientists was bullying and discrimination in their work environment. This was especially common for women, researchers identifying as gay and people of mixed race. Some 60% of the last group, for example, reported that they had experienced discrimination, harassment or bullying at work.

Although the study took place in 2020, most of these issues were present well before the COVID-19 pandemic.

One study, for example, surveyed more than 3,000 physicists and biologists in the United Kingdom, United States, Italy and India, and conducted in-depth interviews with more than 200 scientists. It found that the pandemic only exacerbated issues that were already present[10].

“In our interviews, scientists said that the pandemic was really just the tipping point,” says study co-author Brandon Vaidyanathan, a sociologist at the Catholic University of America in Washington DC. “Things like burnout and emotional exhaustion had been happening beforehand and the pandemic laid them bare. It amplified these effects.”

Chance to change

As evidence of all these problems has grown, scientists are turning towards solutions. But exactly what needs to change is far from clear. “When we looked at the results from the study, we were actually quite upset and felt a little angry — helpless even,” Goel says. “We just felt that it would be extremely difficult to effect change.”

Still, many scientists agree that the first step is for mental health to become a mainstream topic of conversation — a change that might already be afoot. Since Evans and her colleagues published their study in 2018, they and other scientists have been invited to numerous institutions and conferences to talk about the issue. And the most recent decadal survey of US astronomy and astrophysics, in 2021, conducted by the National Academies to set funding priorities every decade, discussed the need to address harassment and discrimination in the community.



Mental health of graduate students sorely overlooked

That’s huge progress, argued Jennifer Wiseman, an astronomer at NASA’s Goddard Space Flight Center in Greenbelt, Maryland, at a meeting of the American Astronomical Society (AAS) in Seattle, Washington, in January. Other programmes are looking to tackle the representation and equity issues

that so often coincide with harassment, abuse and mental illness. For instance, NASA recently reviewed its Hubble Fellowship Program, which provides esteemed postdoctoral positions, in order to diversify its reach. Major grant programmes, such as those of the US Department of Energy, now require a plan that outlines how the proposing team will work against barriers to create and sustain an inclusive work environment. “These examples highlight actions being taken in our astronomical field in response to shifting values that now prioritize inclusivity and the well-being of our communities as central to scientific success,” Wiseman said at the meeting.

Milgram has long viewed mental health as an essential component of a successful researcher’s skill set. Since 2020, the NIH Office of Intramural Training and Education has offered a course called Becoming a Resilient Scientist, with topics that range from impostor syndrome to developing better relationships with mentors. Its 25,000 participants — including undergraduates, graduates, postdocs and medical students recruited from the NIH and various extramural institutions — have reported decreases in anxiety, depression and presenteeism (that is, not fully functioning in the workplace because of stress). In the past two years, Milgram’s team has offered a parallel programme called Raising a Resilient Scientist, which is designed to help faculty members and administrators develop better mentoring skills and work on their personal mental health.

Grassroots initiatives by graduate students and postdocs have led to various events and workshops, even yoga meet-ups. Drives to form PhD student and postdoc unions on US campuses have also sought guarantees of better working conditions. “It is the beginning of a movement that hopefully over generations of academics will result in long-term change,” Evans says. “And I feel like we have all the pieces to see that happen. It just takes time.”

But others argue that science needs larger, systemic changes — such as zero tolerance for abuse. Although institutions and conferences are strengthening value statements, codes of conduct and enforcement, there is still a lot of work to be done, Wiseman said at the AAS meeting. Researchers need to be able to raise concerns without fear of reprisals or prejudice. In Goel’s survey[9], for example, 49% of respondents said that they would not seek support because they would worry about retaliation.

Many argue that there needs to be drastic changes to funding structures. “These models are just really dated,” says Sheila Kanani, the education outreach and diversity officer at the Royal Astronomical Society in London, which surveyed 650 astronomers and physicists in 2020 and found a systemic bullying problem. “The whole system needs a massive overhaul.” As part of that, she and others say that funding should not be entirely based on publications, but rather on a healthy work environment — one that considers the mental health of researchers and the rigour of their work, including ideas that might never land in an academic journal.

Without dramatic change, it’s possible that young researchers will continue to flee the field. “This really is a problem that could potentially affect the future of science if we can’t retain young talent,” Vaidyanathan says. “We owe it to future generations of scientists to create more hospitable workplaces that just allow you to just do the science.”

Nature **617**, 666-668 (2023)

doi: https://doi.org/10.1038/d41586-023-01708-4

References

1. Evans, T. M., Bira, L., Gastelum, J. B., Weiss, L. T. & Vanderford, N. L. *Nature Biotechnol.* **36**, 282–284 (2018).

2. Eaton, N. R. *et al. J. Abnorm. Psychol.* **121**, 282–288 (2012).

3. Clancy, K. B. H., Nelson, R. G., Rutherford, J. N. & Hinde, K. *PLoS ONE* **9**, e102172 (2014).

4. Clancy, K. B. H., Lee, K. M. N., Rodgers, E. M. & Richey, C. *J. Geophys. Res. Planets* **122**, 1610–1623 (2017).

5. Johnson, P. A., Widnall, S. E. & Benya, F. F. (eds). *Sexual Harassment of Women: Climate, Culture, and Consequences in Academic Sciences, Engineering, and Medicine* (National Academies Press, 2018).

6. Wellcome. *What Researchers Think About the Culture They Work In* (Wellcome, 2020).

7. Watts, J. & Robertson, N. *Educ. Res.* **53**, 33–50 (2011).

8. Chirikov, I., Soria, K. M, Horgos, B. & Jones-White, D. *Undergraduate and Graduate Students' Mental Health During the COVID-19 Pandemic* (Univ. California, Berkeley & Univ. Minnesota, 2020).

9. Cactus Foundation. *Joy and Stress Triggers: A Global Survey on Mental Health Among Researchers* (Cactus Communications, 2020).

10. Jacobi, C. J., Varga, P. J. & Vaidyanathan, B. *Front. Psychol.* **13**, 923940 (2022).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01708-4

| Section menu | Main menu |

Books & Arts

- **Is the biggest challenge to scientific thinking science itself?** [22 May 2023]
  • Data torturing, cherry-picking, P-hacking and the invention of tools such as ChatGPT — when it comes to assisting the spread of disinformation science is its own worst enemy, argues a new book.
- **The future of museums and a history of ignorance: Books in brief** [17 Mar 2023]
  • Andrew Robinson reviews five of the best science picks.

- BOOK REVIEW
- 22 May 2023

Is the biggest challenge to scientific thinking science itself?

Data torturing, cherry-picking, P-hacking and the invention of tools such as ChatGPT — when it comes to assisting the spread of disinformation science is its own worst enemy, argues a new book.

- Eric-Jan Wagenmakers



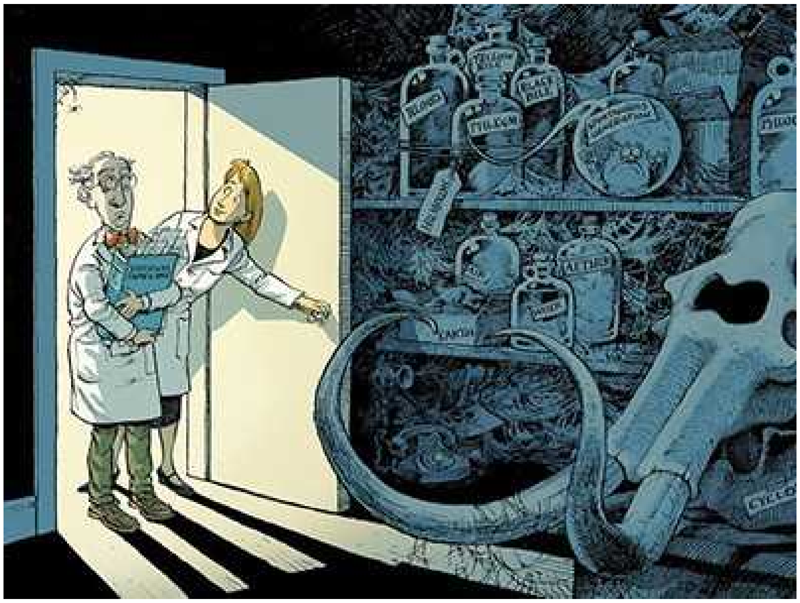
Sloppy science: eating pizza from a small box doesn't necessarily mean you eat less.Credit: Getty

Distrust: Big Data, Data-Torturing, and the Assault on Science *Gary Smith.* Oxford Univ. Press (2023)

Science is under attack. Ironically, the weapons are products of science itself: the propagation of misleading information, the torturing of data to 'prove' claims about anything, the mining of data untroubled by any hypothesis about what you might find. As Gary Smith writes in *Distrust*, "Disinformation is spread by the Internet that scientists created. Data torturing is driven by scientists' insistence on empirical evidence. Data mining is fuelled by the big data and powerful computers that scientists created."

Smith, an economist at Pomona College in Claremont, California, has form in this sort of critique: he wrote the 2018 book *The AI Delusion* and, together with mathematician Jay Cordes, the 2019 book *The 9 Pitfalls of Data Science*. Throughout *Distrust*, he underscores his claims with compelling examples. Take cryptocurrencies, one of his pet peeves. Disinformation and fake trades manipulate their value; data torturing underpins models that supposedly predict their prices; and data-mining creates them in the first place.

He discusses in detail other examples of science being under attack. He sets out how, for example, food-marketing researcher Brian Wansink's claims about dieting — that people eat less if their food comes on a small plate or if their kitchen is painted in neutral earth tones, for instance — were featured in numerous peer-reviewed papers and led to two bestselling books. A classic case of data torturing and sloppy science, the saga known as pizzagate (supposedly the data were largely collected in an Italian diner) eventually led to 18 retractions and numerous expressions of concern about other papers.



Scientists rise up against statistical significance

Then there is the supercomputer IBM Watson, which has data-mining capabilities that would supposedly have revolutionized health care. IBM invested more than US$15 billion on a system that has not yet produced a single peer-reviewed paper but has instead, while employed at the University of Texas MD Anderson Cancer Center in Houston, produced “multiple examples of unsafe and incorrect treatment recommendations”. Don’t even mention former US president Donald Trump and COVID-19, the hydroxychloroquine hoax, conspiracy theories of varying stripes, the fake texts and images created by generative artificial intelligence, claims for the reality of extra-sensory perception, the effectiveness of power posing and so on.

Distrust is a veritable page-turner, and I finished it in a few sittings. On a higher level, it is a call for common sense, for scepticism, for methodological rigour and for epistemic modesty. I suspect most scientists will love it.

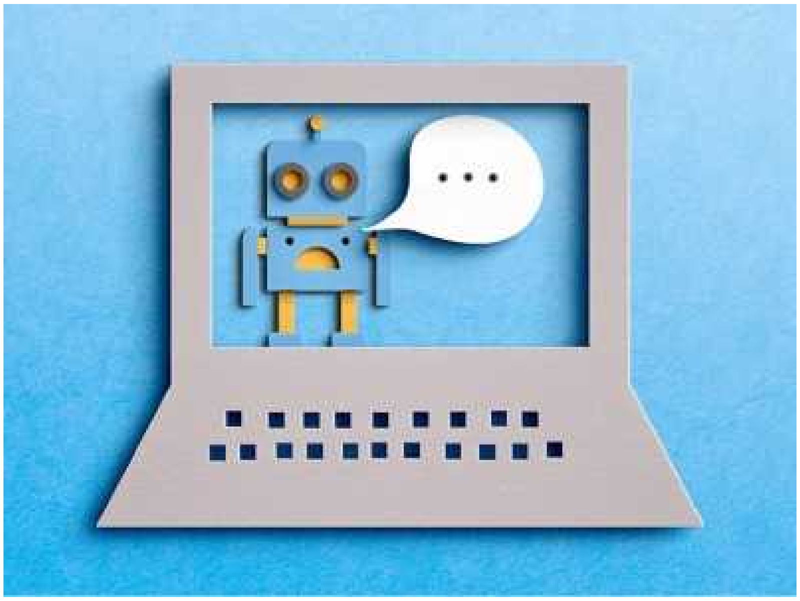
But in places it misses the mark. I found the lack of proper scientific referencing disappointing. I can hardly fault the author, as I have not included explicit references here either, for fear of not fitting the mould of *Nature* book reviews. But a book on disinformation ought to religiously cite its sources for any claim that it makes.



[What 'data thugs' really need]

Other miscues are more notable. Every year, the *British Medical Journal* (*BMJ*) publishes articles in its notorious (and entertaining) Christmas issue that purposely take things to extremes and draw conclusions that are patently ridiculous. Smith seems to take these articles at least semi-seriously. After demonstrating that many of the articles result from cherry-picking and *P*-hacking — torturing out statistically significant effects from data — he discusses a paper in which remote prayer was shown to improve outcomes for hospitalizations that had occurred several years earlier (because the authors were unwilling to assume that "God is limited by a linear time"). At this point, Smith notes, "I read that sentence twice and realized that this was a prank paper." But so are the other *BMJ* papers that Smith critiques.

Distrust also pays little attention to the methodological improvements that scientists have embraced over the past decade to right the ship, or at least to counter data torturing and data mining. Only in the final chapter ('Restoring the Luster of Science') does the author provide a short, superficial discussion of ways to counteract questionable research practices. To my mind, the question of what should be done has a simple answer: academic journals should adopt the practices set out in Level 2 of the Transparency and Openness Promotion (TOP) Guidelines established by the Open Science Foundation in 2014. (Full disclosure: I was part of the committee of scientists that formulated the original guidelines.)



[AI chatbots are coming to search engines — can you trust the results?]

Distrust does mention pre-registration as a possible countermeasure: committing to a specific analysis plan in advance of data collection. But Smith argues that "relatively few journals currently require pre-registration — perhaps because it is so easy to game the system: collect the data, torture or mine the data to obtain interesting results, and then file a pre-plan that does not reveal that the study has already been completed." However, many reputable medical journals effectively require pre-registration for publishing clinical trials. Although scientists can circumvent the rules, doing so would be outright fraud. Researchers who stoop that low might as well just make up the data from scratch.

The book joins a growing chorus saying that schools and universities ought to teach courses in quantitative literacy to counter the wider societal problem of scientific disinformation. French scholar Pierre-Simon de Laplace was already arguing for that in 1814; what *Distrust* often lacks is a prescription for what this might entail. One concrete recommendation is that "statistics courses in all disciplines should include substantial discussion of Bayesian methods". I describe myself as a dedicated Bayesian and would argue that Bayesian methods of statistical inference are the bedrock of all rationality, so I fully support this idea; yet only two pages in the book are devoted to talking about this method.

The broader question is whether any educational initiative would do much good. As a species, humans have always been shockingly biased and gullible. Raw intelligence does not seem to provide much, if any, protection against misinformation. Alongside the entertaining examples of people believing weird things, I would have wanted a discussion of 'why people believe weird things'. Michael Shermer's 1997 book of that title could easily have been the main source for at least one extra chapter. The many examples of bad science in this highly readable, topical book are educational and distressing, but the focus is too much on the disease, and too little on the potential cures. *Distrust* lights a few candles, but mostly curses the darkness.

Nature **617**, 669-670 (2023)

doi: https://doi.org/10.1038/d41586-023-01709-3

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01709-3

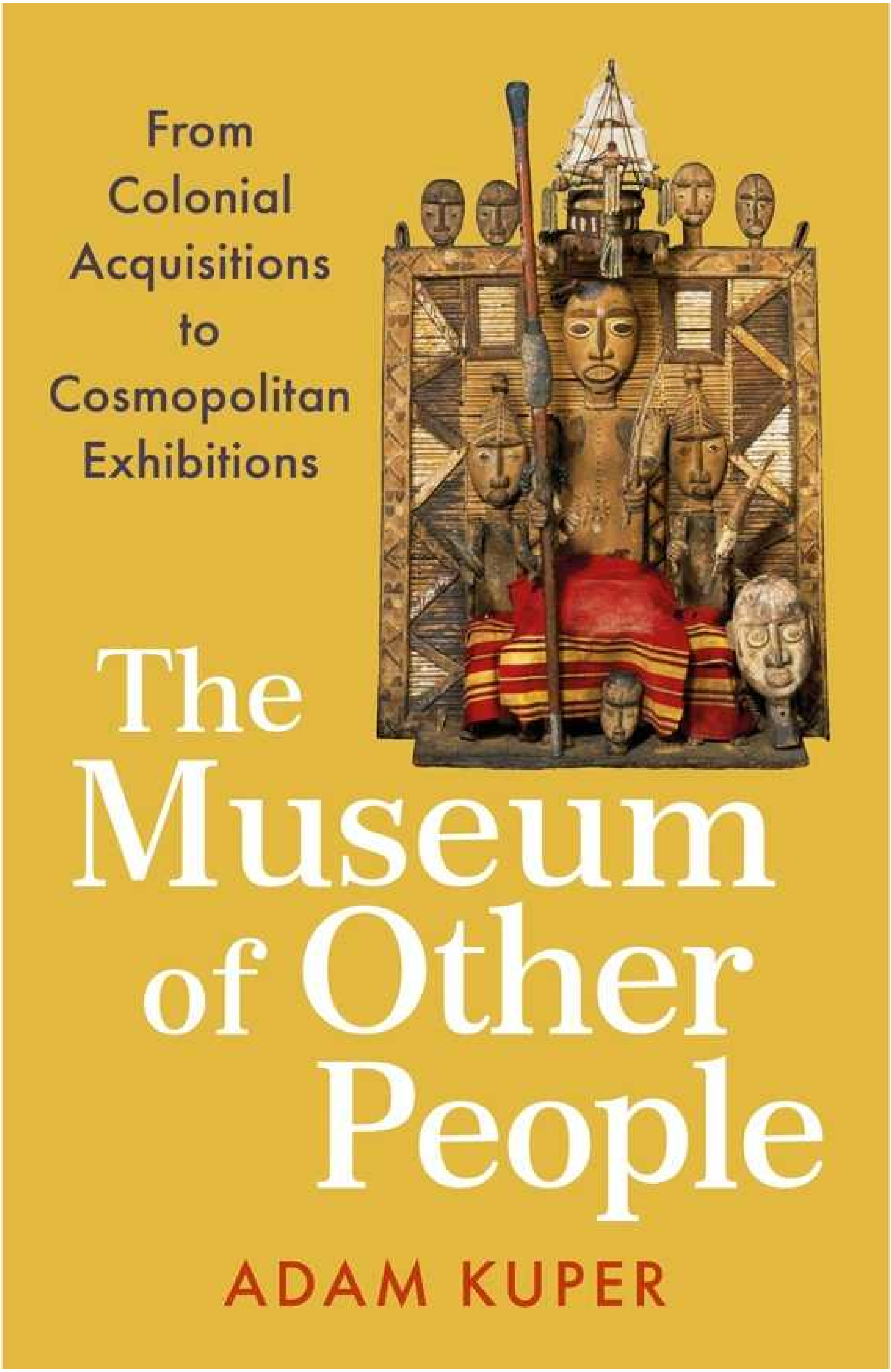
| Section menu | Main menu |

- BOOK REVIEW
- 17 March 2023

The future of museums and a history of ignorance: Books in brief

Andrew Robinson reviews five of the best science picks.

- Andrew Robinson


From
Colonial
Acquisitions
to
Cosmopolitan
Exhibitions
The
Museum
of Other
People
ADAM KUPER

The Museum of Other People

Adam Kuper *Profile* (2023)

This fascinating history by anthropologist Adam Kuper discusses ethnology museums, mainly in Europe and the United States, established during the colonial period. He argues that it is time to turn such institutions into 'cosmopolitan museums' that include challenging perspectives and contrasting points of view — backed by research and scholarship, not "mystical insight" or the "authority of identity". But he also recognizes that "the force is with those who demand the restitution of colonial collections".



Ignorance

a global history

PETER BURKE

'Only Peter Burke could have written such a deliciously knowledgeable history of ignorance.' **David Armitage, Harvard University**

Ignorance

Peter Burke *Yale Univ. Press* (2023)

Having published a study of polymathy in 2020, cultural historian Peter Burke now tackles its opposite: ignorance. Both have long been key to scientific progress. “Thoroughly conscious ignorance is the prelude to every real advance in science,” remarked physicist James Clerk Maxwell in the nineteenth century. Chapters also consider the relevance of ignorance to business, geography, politics, religion and war. Burke argues, with encyclopedic cogency, that we should think of ‘knowledges’ and ‘ignorances’ — plural rather than singular.

Invention and Innovation

A Brief History of Hype and Failure



Vaclav Smil

Author of the *New York Times* bestseller *How the World Really Works*

Invention and Innovation

Vaclav Smil *MIT Press* (2023)

As an environmentalist and energy writer, Vaclav Smil is well placed to analyse the impact of past and promised inventions and innovations. He distinguishes between these concepts: innovation, he says, involves "mastering new materials, products, processes and ideas". He focuses engagingly on three types of "failed" invention: welcomed but then unwelcome (for example, leaded petrol and the pesticide DDT); over-hyped (such as nuclear fission and supersonic flight); and undelivered (including travel by vacuum tube and controlled nuclear fusion).

What is Intergenerational Justice?



Axel Gosseries

What Is Intergenerational Justice?

Axel Gosseries *Polity* (2023)

What duties do we have to future generations, both living and unborn? What does ‘justice’ mean in the intergenerational context? On topics such as climate change, we have to “divide cakes without knowing how many guests will join us and what their tastes will be”, comments political philosopher Axel Gosseries. Despite the emotive subject, his language and tone are academic. Greenhouse gases and extinction are discussed, but not Greenpeace and Extinction Rebellion — even though he accepts the need for “radical changes”.


HEINRICH PÄS
THE
ONE
HOW AN ANCIENT IDEA HOLDS
THE FUTURE OF PHYSICS

The One

Heinrich Päs *Basic* (2023)

Theoretical physicist Heinrich Päs begins his intriguing, controversial book for general readers on the mysteries of quantum physics by asking, how can standing under a star-studded night sky make us feel at once insignificant yet “strangely at home in the universe”? He answers, essentially, that everything — including quantum particles — is part of the same fundamental whole. He is a monist, convinced by a comment by ancient Greek philosopher Heraclitus: “From all things One and from One all things.”

Nature **617**, 670 (2023)

doi: https://doi.org/10.1038/d41586-023-00787-7

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-00787-7

Opinion

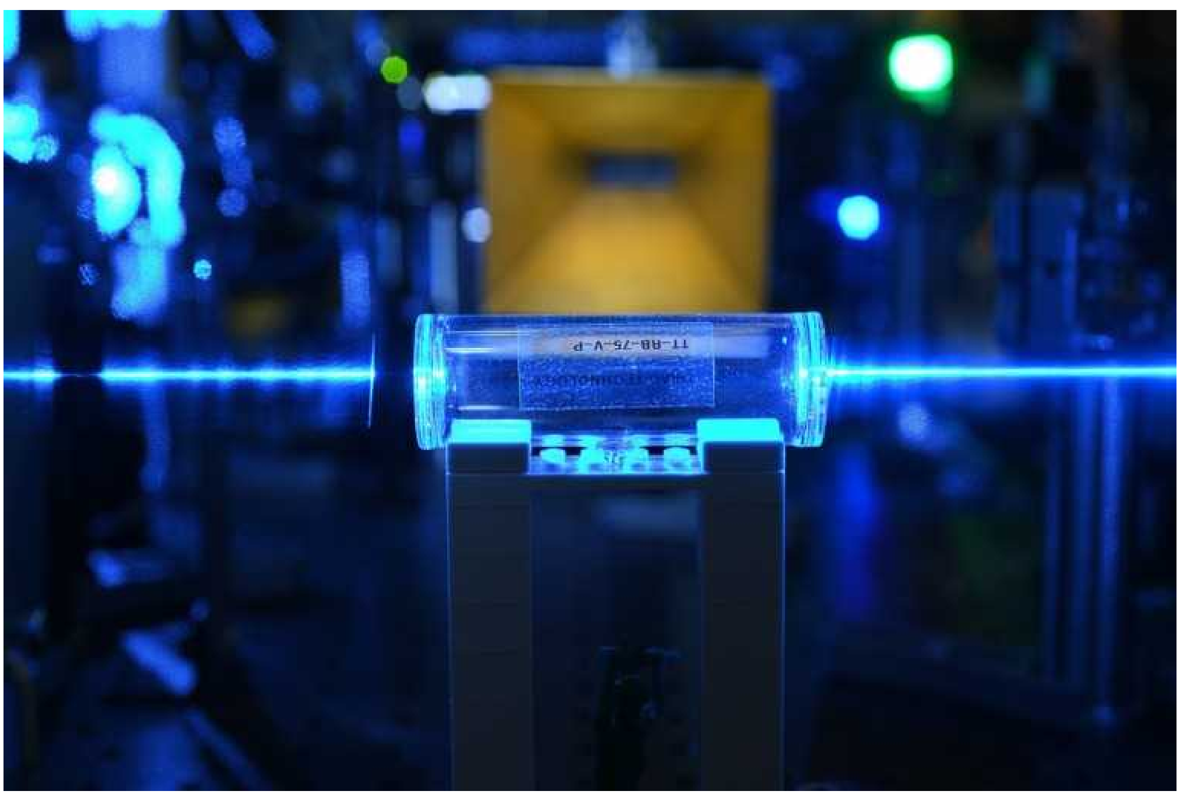
- **Quantum sensors will start a revolution — if we deploy them right** [24 May 2023]
  • From underground exploration to brain science and air-traffic control, the sensing potential of quantum devices is enormous. But they must first get out of the laboratory.

- COMMENT
- 24 May 2023

Quantum sensors will start a revolution — if we deploy them right

From underground exploration to brain science and air-traffic control, the sensing potential of quantum devices is enormous. But they must first get out of the laboratory.

- Kai Bongs ,
- Simon Bennett &
- Anke Lohmann



A quantum sensor developed by the US Army in 2020 can detect communications signals over the entire radio-frequency spectrum. Credit: United States Army

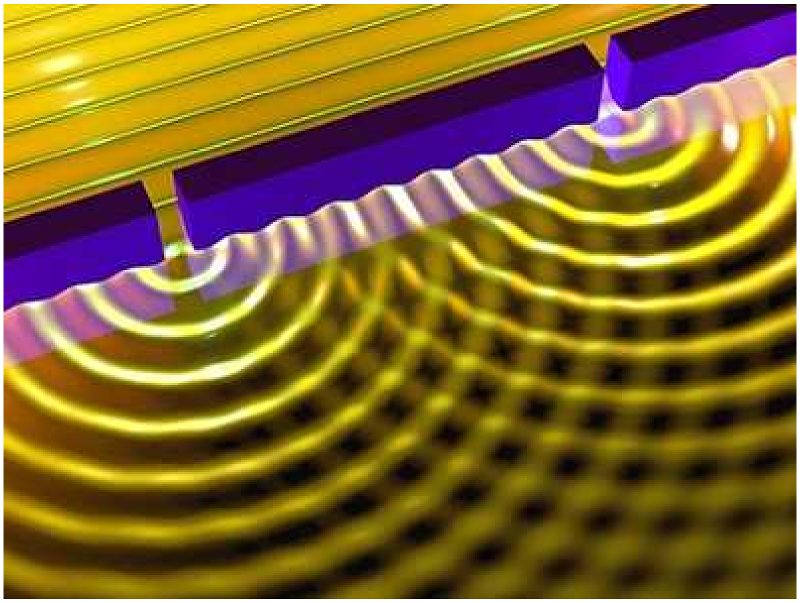
Quantum sensors exploit the fundamental properties of atoms and light to make measurements of the world. The quantum states of particles are extremely sensitive to the environment, which is a virtue for sensing, if problematic for making a quantum computer. Quantum sensors that use particles as probes can quantify acceleration, magnetic fields, rotation, gravity and the passage of time more precisely than can classical devices that are engineered or based on chemical or electrical signals. They can be used to make atomic clocks that are smaller and more accurate, cameras that can see through fog and around corners, and devices for mapping structures underground, among many other potential applications. They stand to transform a multitude of sectors, from energy, land use and transport to health care, finance and security. But their commercial promise needs to be appreciated more.

As researchers developing quantum sensors in the laboratory, we are keen to make governments and industry more aware of the possible benefits — in particular, in improving the safety of national critical infrastructure that relies on sensors, such as air-traffic control systems and water utilities. However, we and others face hurdles in gaining attention and funding to adapt quantum sensors for use in real-world settings.

One challenge is that it is hard to predict exactly how and where emerging technologies will be adopted. The history of physics is full of serendipitous inventions. X-ray generators, for example, were the accidental by-product of experiments to see whether beams of electrons could pass through glass, yet they are now crucial to medicine and airport security. The inventor of the laser, Theodore Maiman, famously described that technology as a "solution seeking a problem".

Another factor is that many people — including business leaders — think quantum technologies are devices of the future, not the present. Unlike quantum computers, which get a lot of press but might be decades away from offering wide commercial advantage, quantum sensors are already in use in the lab. A handful are in commercial use: atomic clocks, for example,

measure the passage of time supremely accurately using high-frequency quantum transitions in atoms. Their accuracy maintains the synchronization of communication and energy networks, and digital radio stations. They are crucial for satellite navigation services such as GPS.



Light waves squeezed through ‘slits in time’

Even so, it took 20 years to move GPS receivers from being specialist devices used by the military, tech-savvy hikers and ships’ captains to providing navigation for smartphones and cars. Now, the quantum community needs to establish similar pathways for realizing the commercial benefit of other types of quantum sensor.

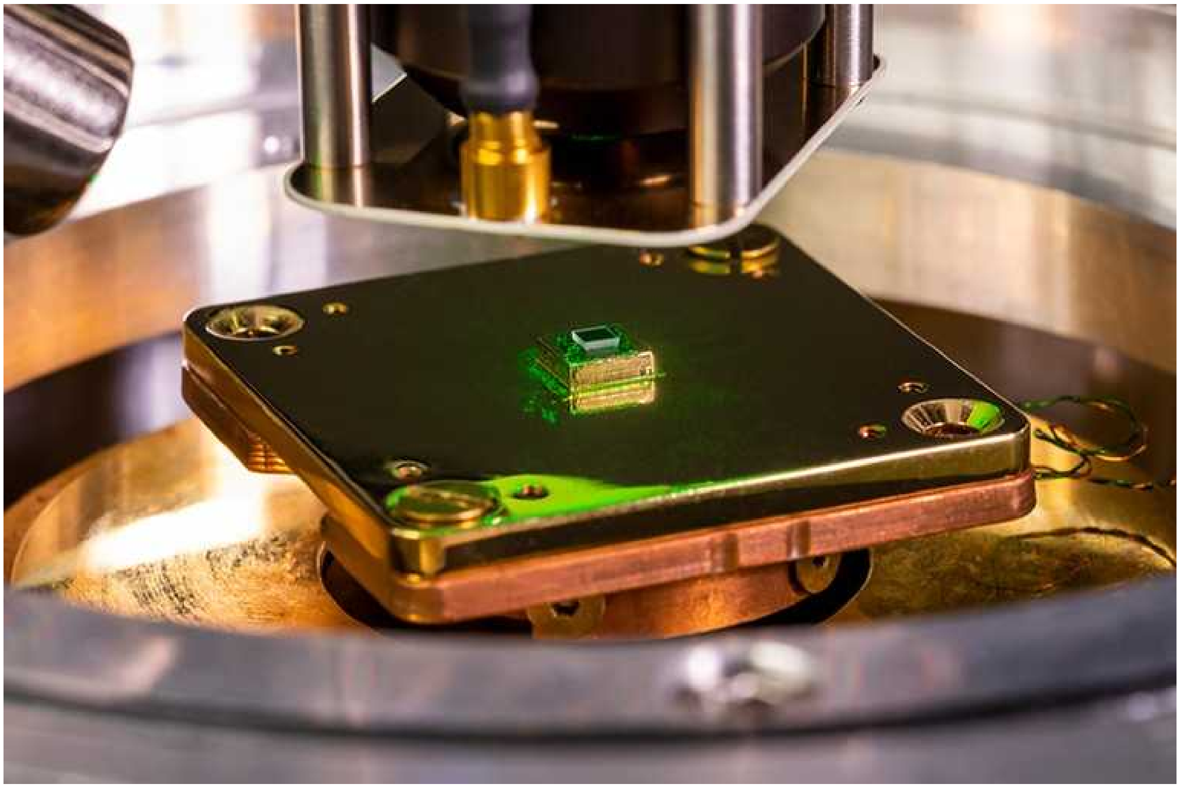
Quantum gravity sensors and quantum gas detectors flown on satellites could collect accurate data on levels of groundwater, carbon dioxide and methane to improve climate modelling. Quantum magnetic sensors can image people’s brain signals in real time[1], and quantum gravimeters can monitor underground water levels and volcanic eruptions[2]. Combinations of quantum sensors that track gravity gradients, magnetic fields and inertial forces 1,000 times more precisely than can classical ones might, for instance, enable reliable navigation in places where satellite signals are jammed or cannot reach, such as remote sites, conflict zones or underwater.

Here we highlight five priorities for commercializing quantum sensors to get them adopted faster.

Make them more robust

Innovators in industry are rarely excited by a lab result that simply proves a concept. They want to know that a device will work reliably for a specific application, and that it will benefit their business's finances. Researchers need to ensure that any sensor bound for the market is robust and reliable, can be manufactured reproducibly and cost-effectively, and is compatible with other systems in use. In practice, this might mean redesigning many aspects of the technology. Each tweak brings fresh challenges.

For example, in our lab at the UK Quantum Technology Hub Sensors and Timing in Birmingham, we have developed a sensor that measures gradients in gravity. In two chambers 1 metre apart, lasers trap rubidium atoms from a vapour and cool them to a standstill[3]. More laser pulses create superpositions of quantum states and read these out for the rubidium atoms in each cloud. Software converts those signals into a gravity gradient measure. By using a single laser beam to manipulate the atoms, this quantum device is 1,000 times less sensitive to noise from vibrations than are conventional gravimeters, and is thus easier to deploy.



A diamond-based quantum sensor can measure magnetic fields at the atomic scale.Credit: David Kelly Crow/de Leon lab/Princeton University

The version we first demonstrated in the lab was the size of a small van, with tables full of optical components and racks of electronic systems and power supplies. It was built from bespoke parts and tuned by hand. Taking this device outside the lab, to sense subterranean tunnels through small changes in local gravity[4], meant making all of the components more robust, smaller and cheaper, as well as improving their performance.

Our physicists and engineers had to find ways to control the laser beam under varying temperatures, contain it in a vacuum to avoid air turbulence and pulse the laser to reduce the impacts of stray magnetic fields. Work is ongoing to operate the device on a moving platform to ease deployment, to increase its sensitivity and bandwidth to speed up mapping, and to reduce its size to that of a backpack so it can be mounted on a drone to survey large areas.



Test quantum mechanics in space — invest US$1 billion

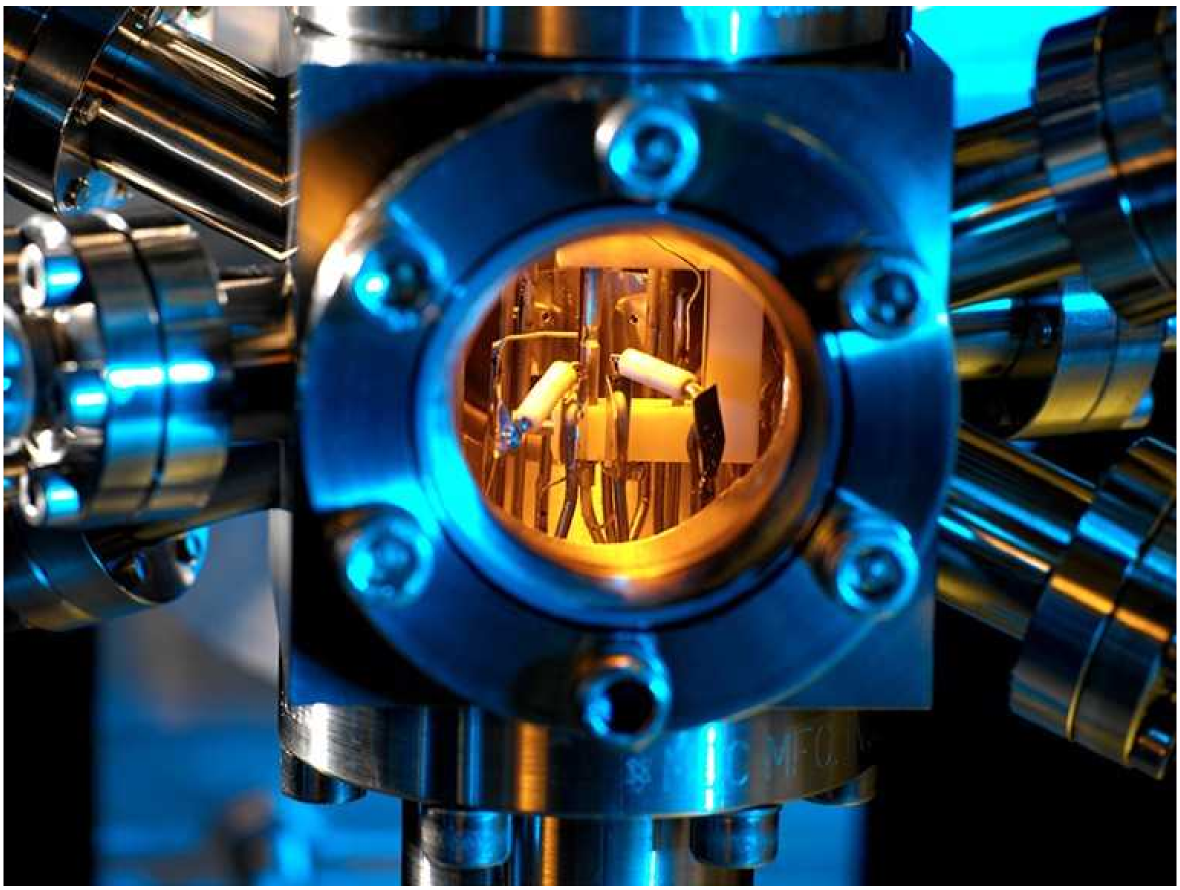
One promising avenue for miniaturization is integrating quantum sensors in photonic microchips[5,6]. These rely on light (photons) rather than the electrons used in conventional microchips, and are fast and energy efficient. Similar technology is found in fibre-optic networks. Quantum sensors could be miniaturized using photonic chips and existing manufacturing processes for micro-electro-mechanical systems (MEMS), which are used in vehicle airbags. The benefit is that they are robust and cope with vibrations better than bulkier optical systems do.

The challenge is integrating all the elements into one system that includes lasers, modulators, waveguides and beam splitters, as well as components such as vapour cells. Further research and investment are needed into new materials, fabrication technologies, device packaging and procedures for testing and validation. Standardization of quantum sensor technologies as low-cost building blocks is also urgently required, mirroring the processes for fibre-optic communication and MEMS sensor technologies.

Work out where we need them

Researchers need to talk to business leaders to determine how quantum sensors can add value across a range of applications. For example, uses for a

gravity sensor are not obvious; few people visualize their surroundings in terms of gravity or the density of materials. But after discussions with more than 100 companies, we concluded that gravity sensors would be excellent for illuminating unknowns in the ground, from the position of forgotten mineshafts to groundwater levels and the distribution of carbon in soils and magma flows. These can, in principle, be seen by classical gravimeters, but ground vibrations make the measurement time infeasibly long, typically 5–10 minutes for a single data point. With quantum gravity gradiometers, such data could be collected in seconds, opening up the potential for gravity cartography[4]. And that's just what we have focused on so far.



An optical clock in which strontium ions oscillate in response to laser light.Credit: Andrew Brookes, National Physical Laboratory/Science Photo Library

Money for applied research and multidisciplinary collaborations between academia and industry is needed to validate these ideas. In our case, the next

step involves geophysics research using such gravimeters to improve understanding of how water flows and accumulates underground. This information could be used to refine flood models, for example. Civil-engineering research is also required on how best to detect leakage in water pipes using such sensors. Broader technological and economic considerations will determine how this approach can be used most effectively in water management.

Companies should start thinking about new business models, such as offering underground mapping services to farmers to help reduce water use in irrigation. Engaging in pilot projects would put businesses in a good position to capitalize on market disruption, rather than being caught out by it.

Integrate them in current systems

Any sensor must be plugged into a bigger system to reap its benefits. For example, an inertial sensor — one that detects movement — is relatively useless on its own. But when integrated with electronics, software and a display in a smartphone, it can provide health information on step counts and calories burnt by the user.

Similarly, quantum accelerometers and sensors of rotation, gravity and magnetic field can be combined into position, navigation and timing systems for subsea and underground use. For this application, quantum sensors offer reduced bias, better precision and more stability than do their classical counterparts, allowing navigation with metre-level accuracy without having to use global satellite systems such as GPS. This capability would enable exploration of resources on the seabed, as well as securing and maintaining pipelines, cables and foundations of offshore wind farms and oil rigs, for example.

However, it remains fiendishly challenging to integrate quantum sensors into a full-blown navigation system. Constructing an inertial measurement unit alone requires three accelerometers, one for each spatial dimension, and three rotation sensors, one for each rotational degree of freedom, arranged at perfect right angles and all linked with a clock. If fitted on a vehicle or

submarine, such a navigation system would need to compensate for small changes in local gravity and other forces induced by Earth's rotation. The whole thing would need to be calibrated, which is hard to achieve at the high level of precision needed.



[Quantum computing's reproducibility crisis: Majorana fermions]

Gravity and magnetic sensors would be needed for mapping these fields along the trajectory of the vehicle, as well as a computer control system with specialist software. Databases of field readings would need to be developed for comparison against the recorded gravity and magnetic traces, to allow absolute position fixes to deal with unavoidable long-term drifts.

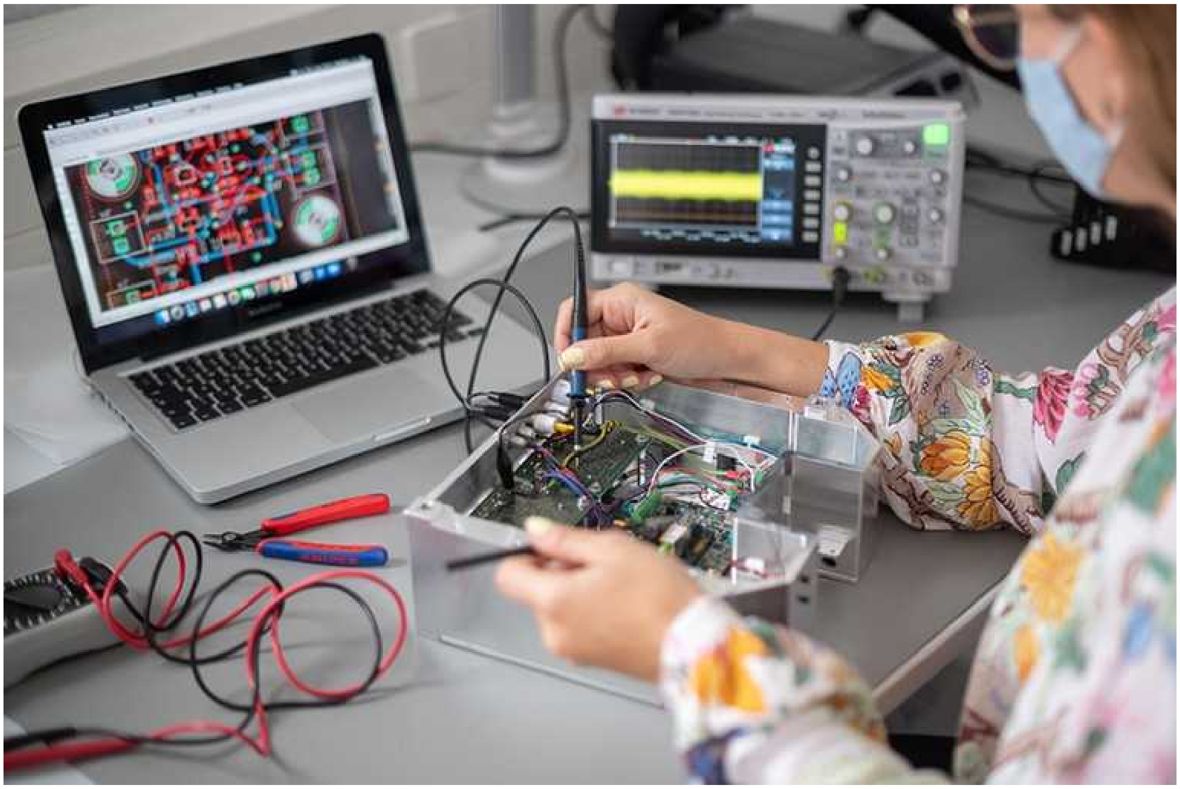
Researchers also need to consider in detail how quantum sensor systems might be linked to national and international infrastructure networks. For example, communication networks could be revolutionized with the next generation of quantum clocks, 'optical clocks', which could be 1,000-fold more precise than the time provided by current satellite navigation systems. This might enable new modes of ultra-fast broadband, for example, which squeeze more data packets into channel bandwidth and use less energy to transmit each bit of data. Similarly, quantum sensors capable of detecting hydrogen could speed up the energy transition from natural gas to hydrogen fuels, because they could detect leaks and safeguard infrastructure to enable secure roll-out of this potentially highly explosive fuel.

Whereas academic researchers can develop sensors with the right properties, industry needs to lead this systems integration stage. Existing academic funding streams are too small to support such collaborative research. Substantial long-term research and development contracts with industry are needed to make this happen. For instance, in the 2000s, funding from the US Defense Advanced Research Projects Agency helped to create the chip-scale atomic clock within a decade, through a dedicated development programme involving academia and industry.

Establish data needs

Raw data from a sensor needs to be transformed into information that is useful for a specific task. For example, although a quantum magnetic-field sensor can detect tiny fields associated with patterns of neural activity in the brain, 3D visualizations of brain activity require an array of such sensors, and algorithms and graphical representations to display them in ways a physician can interpret.

Development of such systems is under way[1] and could revolutionize understanding of brain conditions. Real-time mapping (scans 100 times a second, for example) and analysis of brain responses to visual or sensual stimuli, even while a person is moving, might replace current techniques for diagnosing brain disorders based on patient questionnaires. It could also allow physicians to assess the efficacy of drugs for brain conditions on an individual basis.



A researcher at German firm Q.ANT checks a quantum sensor intended for industrial use.Credit: Sebastian Gollnow/dpa via Alamy

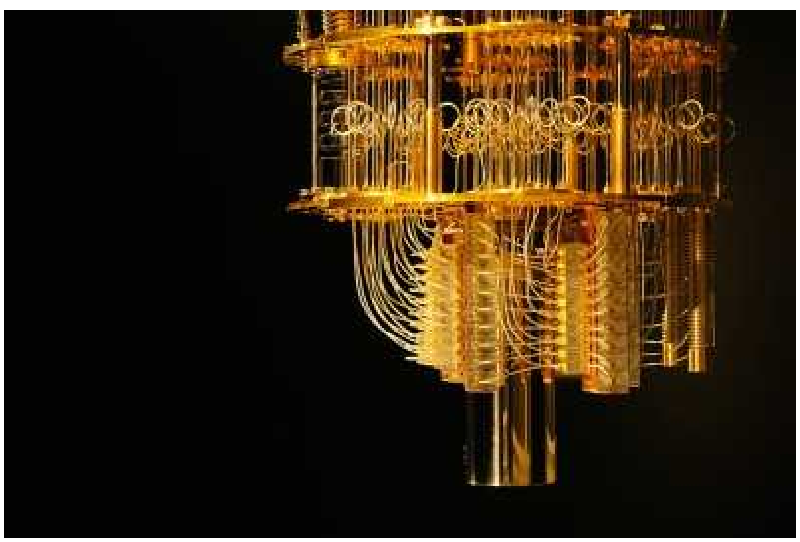
Similarly, advanced analytics are needed to extract 3D underground images from gravity surveys, where it remains challenging to determine how deep sensed objects are. Banks of radars driven by quantum oscillators need to be networked to show detailed images instead of dots on a radar screen, such as would be needed to classify and distinguish drones from birds flying over a city. Big data techniques must be deployed to harvest all of this information, enabling the monitoring of tens of thousands of delivery drones in cities, for instance.

Perhaps the greatest data challenge in terms of time and effort is to create 'training' data sets through trials. Researchers need to conduct large-scale medical trials to find biomarkers for brain conditions, collect data from networks of gravimeters to understand underground water and other assets, and source radar data through sensor networks across cities. We encourage governments to fund such programmes to seed future ventures.

Collaborate to innovate

Although many countries have begun coordinated efforts to develop the base level of quantum technologies, there is still a scattered approach to adoption and exploitation. With many groups working in isolation, tackling the research challenges we outline would take decades. To speed things up, a strategy for coordinating research projects on quantum sensors is needed.

At the research end of the pipeline, some nations, including Germany, Japan, the Netherlands, the United Kingdom and the United States, have set up hubs and large projects to align academic and national needs in quantum technologies by bundling expertise and providing portals for interactions with industry and other partners. Yet, generally, sensors are not getting the attention they deserve in national quantum tech initiatives, with a few exceptions, such as QuantumBW, an initiative by the German state of Baden Württemberg, which explicitly focuses on quantum sensing.



Underdog technologies gain ground in quantum-computing race

Governments need to introduce policies and regulation to support innovation in quantum sensors, with one focus being enhancements to the management and security of critical national infrastructure. For example, a 2020 presidential order requires US national aviation authorities to become independent of global navigation satellite systems timing by 2025. This would ensure air-traffic control systems keep working even if those systems

fail by accident or through hostile intervention. It is still too early to determine the impact, but the order has set the boundary conditions for the emergence of business ideas related to timing technologies.

Similar approaches in communications, water-resource management and medicine might encourage the uptake of quantum sensors in those sectors to make them more resilient — by having independent timing and navigation or more detailed data.

Initiatives are also needed to bring companies, from component manufacturers to system integrators, together with academics to help find business solutions, rather than simply come up with the technology and then quickly scale up production in the hope that there will be a market. One promising effort is the National Accelerator for Quantum Sensors in the United Kingdom. Launched in 2022 and still to be fully funded, this accelerator involves three corporate giants with a global reach (BAE Systems, BP and BT) and is committed to bringing in dozens more companies. Although initiatives in other countries target quantum technologies in general — such as QED-C in the United States — the UK programme is unique in that it focuses on sensors.

To conclude, a long-term, industry-led approach for quantum sensor innovation is urgently needed. The physics of quantum sensors can deliver the performance, but the question is: who will lead the world in delivering the benefits?

Nature **617**, 672-675 (2023)

doi: https://doi.org/10.1038/d41586-023-01663-0

References

1. Boto, E. *et al. Nature* **555**, 657–661 (2018).
2. Antoni-Micollier, L. *et al. Geophys. Res. Lett.* **49**, e2022GL097814 (2022).
3. Phillips, W. D. *Rev. Mod. Phys.* **70**, 721 (1998).

4. Stray, B. *et al. Nature* **602**, 590–594 (2022).

5. Newman, Z. L. *et al. Optica* **6**, 680–685 (2019).

6. Dyer, S. *et al. Phys. Rev. Appl.* **19**, 044015 (2023).

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01663-0

| Section menu | Main menu |

Work

- **How mixing academia and industry opens doors in graduate school and beyond** [22 May 2023]
  • A growing number of PhD programmes and postdoc positions combine academic questions and industry resources.
- **Ranching sea cucumbers to repair the oceans — and improve human health** [22 May 2023]
  • Libin Zhang aims to restore sea-cucumber populations to rebalance ocean ecosystems and so that people can use their beneficial compounds.

- CAREER FEATURE
- 22 May 2023

How mixing academia and industry opens doors in graduate school and beyond

A growing number of PhD programmes and postdoc positions combine academic questions and industry resources.

- Freda Kreier

Find a new job



Patience Chihomvu pitched her postdoctoral project proposal to pharmaceutical firm AstraZeneca in Gothenburg, Sweden, in 2022.Credit: Andy Dunbar/The Burstproof Bubble

Completing a postdoc at a company wasn't always on the radar for Patience Chihomvu. But in 2021, the biotechnologist and recent PhD graduate was facing an all-too-common issue: how to fund her research.

Chihomvu had accepted a postdoc position at the University of the Witwatersrand in Johannesburg, South Africa. Her plan was to search for potential therapeutic compounds in the traditional medicinal plants that her father, a resident of rural Zimbabwe, had taken while recovering from COVID-19.

In doing so, Chihomvu hoped to find new treatments for asthma and other conditions such as long COVID. But the academic laboratory she intended to work in didn't have the means to fully support her work. "There are a lot of hurdles you need to cross, especially here in Africa," she explains. "The competition is very, very tough."

In 2022, Chihomvu applied to join the AstraZeneca R&D Postdoctoral Challenge, a competition, hosted by the pharmaceutical firm, that promised winners at least two years of project funding and support. Winners retain ownership of their original idea, but any development of intellectual property during their postdoc could be jointly owned with AstraZeneca. Later that year, she was one of six finalists flown to the company's research centre in Gothenburg, Sweden, to pitch her research in person to a panel of judges from various research institutions.



The career insights I'm bringing back to academia after a year at Google

During their visit, the finalists toured AstraZeneca's facility. After years in academia, having to wait for supplies to be delivered to her PhD lab at Vaal University of Technology in Vanderbijlpark, South Africa, "I was blown away by everything", she recalls. "I felt like I'd died and gone to science heaven."

In the end, AstraZeneca funded all six finalists' research. A chance at science heaven might seem rare, but this kind of model, which blends academic ideas with industrial resources, isn't unique. AstraZeneca's contest is one example of a temporary, industry-based position that allows academic scientists to get a taste of working at a company, which is, after all, the type of employer that many PhD holders will ultimately have.

This glimpse of industry isn't just restricted to the post-PhD stage: around the world, institutions are offering PhD programmes that — through internships or company-funded projects — straddle the line between academia and industry. The potential of these industry-based PhD and postdoc positions to bridge the gap between university research and commercial products has led some countries, such as Australia, to invest heavily in their development.

Those who participate in these hybrid positions gain new ways of thinking and career options beyond academia. "This is a life-changing opportunity for me," says Chihomvu.

Switching destinations

PhD and postdoc positions that combine academia and industry aren't new. But their prevalence varies widely between disciplines, institutions and nations.

At German technical universities, graduate programmes in which companies help to fund some or all of a student's research are "very common", says Hans-Joachim Bungartz, a computer scientist and dean of computation, information and technology at the Technical University of Munich in Germany. These arrangements are a by-product of these institutions'

mandates, most of which date back to the nineteenth century, to produce an engineering workforce rather than "to produce Nobel prize" wins, he says. Today, the programmes remain popular among PhD candidates as a way of getting acquainted with future employers — and are a way for companies to recruit talent while expanding their research capabilities.

However, it's relatively rare for PhD programmes in the United States — the country that produces the most PhDs in the world — to include any elements of industry, says Fan Yang, a bioengineer and co-director of the US National Institutes of Health (NIH) biotechnology training programme at Stanford University in California.

In the United States, earning a PhD in science, technology, engineering and mathematics (STEM) has historically been viewed as a launch pad into academia. As such, PhD programmes mostly train students in the skills they will need for a lifetime in an academic lab — including running experiments, securing grants and writing papers.



Why industry internships can be your 'golden ticket' to a prosperous career

But this "monk-like existence does not prepare you for the careers that PhD graduates are actually having", says David Madigan, provost and senior vice-president for academic affairs at Northeastern University in Boston,

Massachusetts. The percentage of US PhDs who move on to academic positions has been shrinking drastically.

In the 1990s and early 2000s, at least 49% of US PhD graduates had academic positions lined up after their thesis defence. In 2020, the percentage of PhD graduates headed into industry or business (40%) surpassed the figure for academia (39.6%) for the first time in US history. And as of 2021 — the most recent year for which survey data are available — just 35.9% of recent PhD graduates had academic jobs lined up after graduation, compared with 43.2% heading into industry or business (others were bound for jobs in government or the non-profit sector).

This changing landscape has forced some universities to reconsider what skills and experiences higher education should offer. As a result, a handful of US institutions are experimenting with programmes that introduce students to industry earlier in their careers — often through internships.

The industry internship

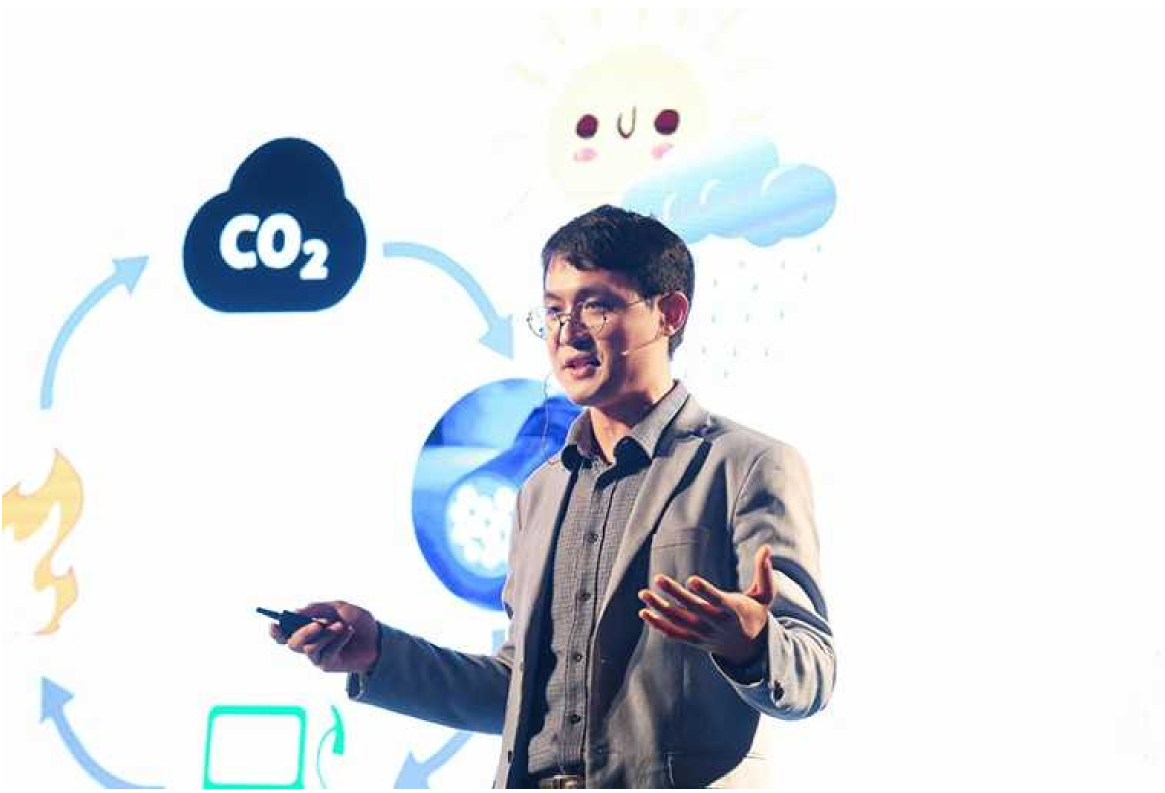
If there's one thing Northeastern University is known for, it's that its students are required to spend entire semesters not going to class. Undergraduates alternate semesters attending university and working full time, the idea being that this real-world experience gives Northeastern graduates an edge in the job market.

Although the university's master's and PhD students are not required to spend time working outside academia, they can opt to do so. In 2019, the school launched a programme to help doctoral students to get internships at government agencies, non-profit bodies or companies while earning their degree.

This voluntary programme was born, in part, out of the "growing realization that a PhD is no longer training to be a professor", says Madigan, who helps to run the initiative. Internships can provide a sneak peek at the private sector, the various job opportunities there and how the skills gained in graduate school can be leveraged in the job market, he says.

Spending time outside academia provides early-career researchers with other benefits, too. In 2022, chemical engineer Xiangkun (Elvis) Cao interviewed five scientists who'd sought out company internships during their time in graduate school (Z. F. M. Burton & X. E. Cao *Matter* **5**, 4100–4104; 2022).

Graduate students reported that internships helped them establish ties to industry, gain new skills and kept them motivated. As one interviewee told Cao: "An internship may be one of the most useful classes you take in grad school."



Elvis Cao says all PhD students can benefit from doing industry internships during their training.Credit: Xu Liu

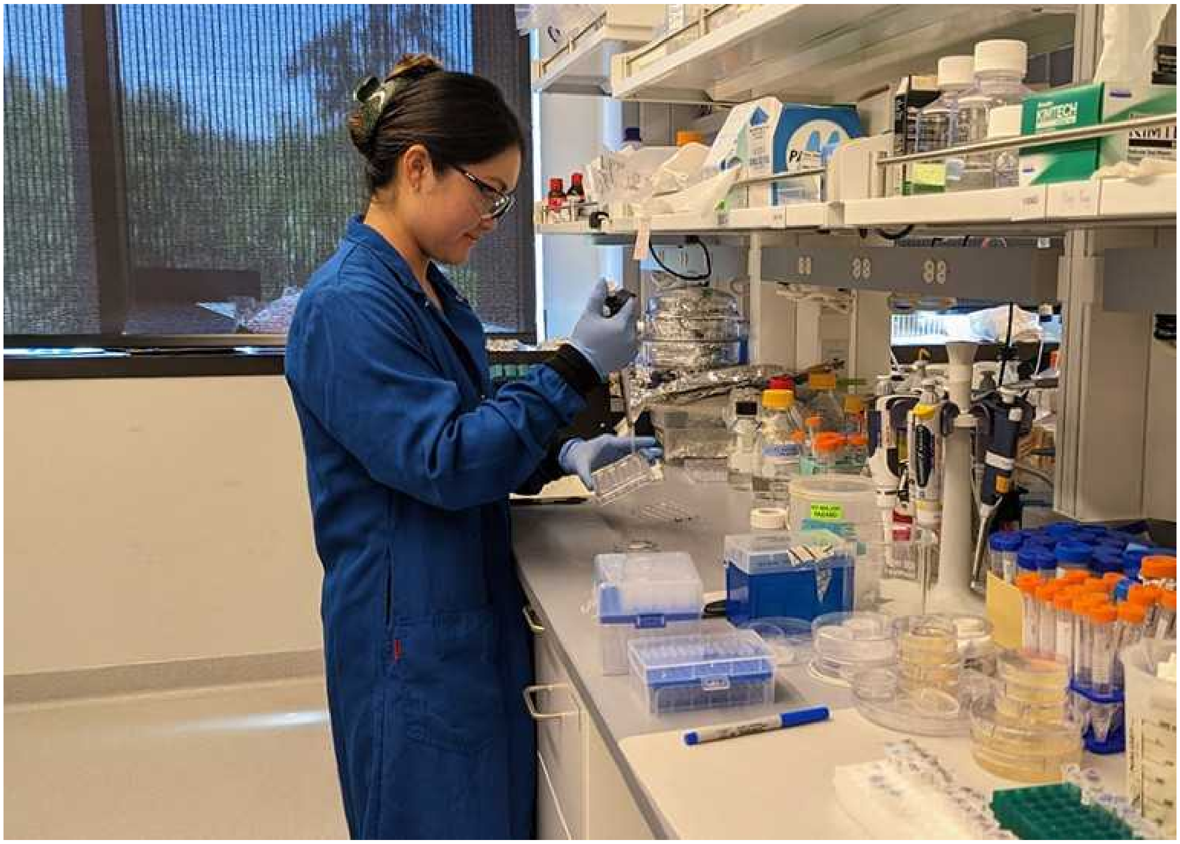
In some fields of study — such as computer science, where most students expect to get jobs outside academia — doing an internship while earning a degree is the norm. But this isn't the case for other disciplines. In addition, convincing a PhD supervisor to allow students to spend time away from their thesis research, potentially for months at a stretch, can be challenging in fields that don't have a culture of internships, says Cao, now a postdoc at the Massachusetts Institute of Technology in Cambridge.

Creating a formal internship programme can make it easier for graduate students to persuade their advisers to let them take leave. Or better yet, the internship can be mandatory: Stanford's NIH biotechnology PhD programme requires each student to complete a three-to six-month internship.

The programme directors chose the requirement because most of their students plan to have a career outside academia, says Yang. Work experience at technology start-ups or major companies in nearby Silicon Valley makes students more attractive to future employers, she says.

"It really helps them have a smoother transition to the real world," adds Yang. It's also a big draw for prospective students. Melissa Nakamoto, a fifth-year PhD student in the programme, says that the internship requirement was a big reason why she chose to apply.

"I've always thought that I would want to work in industry, but it's hard to know for sure when I've really only worked in academic labs," she says. Nakamoto worried that a conventional PhD programme wouldn't give her the time and space to test the waters at an industry internship.



Melissa Nakamoto chose the biotechnology PhD programme at Stanford University because it requires an industry internship.Credit: Aofei Liu

She spent last summer working on a microscopy project for the biotechnology company Amgen in San Francisco, California. Although some of her peers came out of their internships feeling like a small cog in a large money-making machine, her experience felt similar to academia, she says.

“I was still able to do discovery science, but it was more streamlined because I had my own hood, my own microscope and all the chemicals I needed right there without having to worry about whether I had the budget for it,” she adds.

The pool of US PhD programmes that include internships remains small, but more could be on the horizon. Sara Wadia Fascetti, vice provost for the PhD network at Northeastern, says that since the university launched its internship programme in 2019, around half a dozen institutions have asked for advice on how to set up their own system.

Nakamoto thinks it’s too soon to tell whether her internship will guarantee a job after she defends her thesis. “I do think that adding those skills on my résumé makes me a lot more attractive as a candidate,” she says. “Doing an internship opened my eyes to what’s out there.”

On company time

Of course, internships aren’t the only route for people to integrate elements of industry into their academic research. Some early-career researchers have company involvement baked into their PhD from the start.

That’s the case for a handful of PhD students at CERN, Europe’s particle physics laboratory near Geneva, Switzerland. In 2001, CERN created an initiative called openlab to increase collaborations with industry. In recent years, openlab has trained PhD students who split their time between coursework, research and working with an external company.

These students earn their doctorate through a partner university but spend up to 36 months physically working at CERN, while also collaborating with companies.

One of them is Kristina Jarůšková. For her PhD research at the Czech Technical University in Prague, she works with technology companies IBM and Intel to simulate particle interactions inside CERN’s Large Hadron Collider. “Right now, private companies have the most resources in terms of data, people and computing,” she says. She has the opportunity to access these vast resources while also working with leading experts on deep learning and artificial intelligence — most of whom are scooped up by tech giants.

Getting these projects up and running wasn’t smooth sailing, says Alberto di Meglio, who was head of openlab until earlier this year. Firms were initially hesitant to share anything with students that might count as a trade secret.

“For some students, the experience at the company was not that great because they were parked just outside of the main lab and fed the minimum amount of information,” says di Meglio.

To work around these issues, openlab had students sign a non-disclosure agreement in advance. Today, industry-partnered PhDs are well established in the information-technology side of CERN, says di Meglio.

For companies, sponsoring these joint projects can yield the benefit of eventually hiring CERN-trained students who are well versed in the firms' culture and technology, for the small price of US$50,000–60,000 a year, he says. Both the lab and students have also reaped the rewards of these collaborations — including the development of CERN's specialized software. The programme has been so successful in terms of research quality and student motivation "that we decided we would do this kind of PhD [any time] the opportunity arises", adds di Meglio.



[Leaving academia for industry? Here's how to handle salary negotiations]

And some countries, such as Australia, want to make industry-partnered PhDs even more common. Many Australian PhD programmes have been offering connections between industry and academia — often with students agreeing ahead of time to work on a project that interests a company. But these types of partnership made up only around 10% of the total PhD cohort, historically. In 2022, the Australian government announced that it would invest Aus$296 million (US$198 million) to support 1,800 new industry-connected PhDs in the following 10 years.

The move is part of a larger push to bridge the 'valley of death' between academic research and business, where promising basic research often fails

to get picked up and commercialized by private companies, says Deborah Terry, vice-chancellor and president of the University of Queensland in Brisbane, Australia. Terry worked on the government report that outlined how university research could benefit the economy, including through industry PhDs.

Industry PhDs can also help students to chart a career course into the private sector after graduate school, says Zainab Mustafa, a process engineer at the engineering and consulting company Jacobs in Sydney, Australia.

In 2018, Mustafa was one of the first students to do a chemical-engineering industry-linked PhD at the University of New South Wales in Sydney. Mustafa hadn't originally intended to seek a PhD. But she was intrigued by the chance to work in a university lab and also have contacts at the chemical company DuPont and at the Commonwealth Scientific and Industrial Research Organisation (CSIRO), the Australian government's science research agency.



After her industry-linked PhD, Zainab Mustafa landed an engineering consultancy job.Credit: Zainab Mustafa

Working out how to straddle all three worlds wasn’t easy. Mustafa spent the first year of her PhD focusing mainly on industry projects, before realizing that she would need to run experiments that she could submit for a PhD thesis. Looking back, Mustafa says a more structured programme might have helped. “It was still a key learning experience,” she says, adding that she received a job offer from Jacobs about six months before the end of her PhD.

Compared with PhD students working exclusively at universities, Mustafa says, “it was a lot easier for me to get employed because I had a working relationship with industry already”.

This distinguishes industry PhDs and postdocs from more-conventional academic positions. “A lot of PhD graduates feel that their research isn’t going to count towards what industry wants,” says Mustafa. “But I was able to showcase my research and say, ‘I worked with industry to solve their problems in this PhD role’.”

Around half of PhD students in Australia aspire to a career in academia, says Terry. For many in the other half, an industry-linked PhD, internship or postdoc offers what might be the most efficient way to gain skills and experience that translate to the private sector.

That’s the sense Chihomvu has about her AstraZeneca postdoc, which begins next month. Ultimately, Chihomvu plans to stay in industry because it is the best way to make an impact on society, she says. “I cannot wait to start.”

Nature **617**, 851-853 (2023)

doi: https://doi.org/10.1038/d41586-023-01710-w

This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01710-w

- WHERE I WORK
- 22 May 2023

Ranching sea cucumbers to repair the oceans — and improve human health

Libin Zhang aims to restore sea-cucumber populations to rebalance ocean ecosystems and so that people can use their beneficial compounds.

- James Mitchell Crow



Libin Zhang is executive deputy director of the Engineering Laboratory for Marine Ranching in the Institute of Oceanology of the Chinese Academy of

Sciences. Credit: Yue Wu for *Nature*

I study sea cucumbers, an economically important species in China that is valued for its nutritional and medicinal properties, and is important for the health of the ocean. It cleans the sea floor by feeding on detritus. Thanks to climate change and overfishing, sea-cucumber numbers are falling.

At the Engineering Laboratory for Marine Ranching (ELMR), we study the animals to maximize their contributions to marine ranching — a form of sustainable aquaculture that restocks valuable species, such as the sea cucumber and the *Sebastes schlegelii* rockfish. First we transplant seagrass, and then we release sea cucumbers to clean degraded habitats. Next, we release fish, clams and other creatures to build a balanced ecosystem. The productivity is much higher than what would result from restoring one species.

In this picture, taken in the ELMR, which is in the Institute of Oceanology of the Chinese Academy of Sciences in Qingdao, I am looking at *Apostichopus japonicus*, the most important sea cucumber in northern China. In our tanks, we control temperature and pH to mimic ocean conditions. We study how sea cucumbers move, eat, interact and reproduce, which has helped us to design artificial reefs for the sea cucumbers we release.

A. japonicus has some powers that are almost magical. When the water rises above 23 °C in early June, this species begins aestivation, a summer hibernation. It loses its intestine, and regenerates the entire organ when it wakes up in October.

Sometimes, we find white or purple examples of *A. japonicus*, although it is usually green. The purple one is beautiful and contains beneficial compounds similar to those found in red wine, including astaxanthin, which has potent antioxidant and anti-inflammatory properties.

I hope that sea cucumbers can help to repair the ocean, and to repair the body.

Nature **617**, 856 (2023)

doi: https://doi.org/10.1038/d41586-023-01711-9

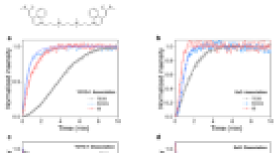
This article was downloaded by **calibre** from https://www.nature.com/articles/d41586-023-01711-9

| Section menu | Main menu |

Research

- **Ångström-resolution fluorescence microscopy** [24 May 2023]
  • The authors introduce a single-molecule DNA-barcoding method, resolution enhancement by sequential imaging, that improves the resolution of fluorescence microscopy down to the Ångström scale using off-the-shelf fluorescence microscopy hardware and reagents.
- **Flexible solar cells based on foldable silicon wafers with blunted edges** [24 May 2023]
  • Modules of foldable crystalline silicon solar cells retain their power-conversion efficiency after being subjected to bending stress or exposure to air-flow simulations of a violent storm.
- **Catalytic disconnection of C–O bonds in epoxy resins and composites** [26 Apr 2023]
  • The authors report a transition-metal-catalysed protocol for recovery of polymer building block bisphenol A and intact fibres from epoxy composites, demonstrating that chemical recycling approaches for thermoset epoxy resins and composites are achievable.
- **A weakly structured stem for human origins in Africa** [17 May 2023]
  • An analysis of models of human populations in Africa, using some newly sequenced genomes, finds that human origins in the continent can best be described by a weakly structured stem model.
- **GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19** [17 May 2023]
  • An analysis of 24,202 critical cases of COVID-19 identifies potentially druggable targets in inflammatory signalling (JAK1), monocyte–macrophage activation and endothelial permeability (PDE4A), immunometabolism (SLC2A5 and AK5), and host factors required for viral entry and replication (TMPRSS2 and RAB2A).
- **Multisensory learning binds neurons into a cross-modal memory engram** [26 Apr 2023]
  • Multisensory learning improves subsequent memory performance, even for individual sensory cues, in Drosophila.
- **Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma** [17 May 2023]
  • Tumour-infiltrating lymphocytes from glioblastoma can recognize bacterial and gut microbial peptides.

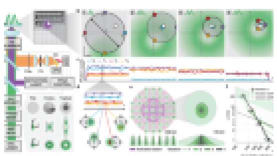
- **[Structural basis of catalytic activation in human splicing]**
  [10 May 2023]
  • Cryogenic electron microscopy images of a spliceosome complex undergoing catalytic activation provide mechanistic insight into how the two ATP-dependent RNA helicases involved in this process, PRP2 and Aquarius, work together.



Three-dimensional super-resolution fluorescence imaging of DNA

27 July 2020

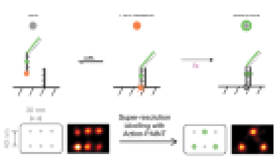
Sevim Yardimci, Daniel R. Burnham, … Hasan Yardimci



MINFLUX nanoscopy delivers 3D multicolor nanometer resolution in cells

13 January 2020

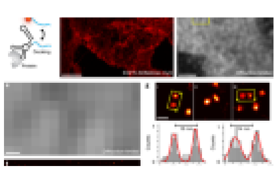
Klaus C. Gwosch, Jasmin K. Pape, … Stefan W. Hell



Super-resolution labelling with Action-PAINT

16 September 2019

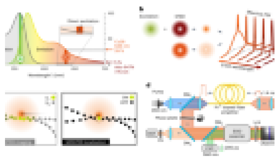
Ninning Liu, Mingjie Dai, … Peng Yin



Modified aptamers enable quantitative sub-10-nm cellular DNA-PAINT imaging

20 August 2018

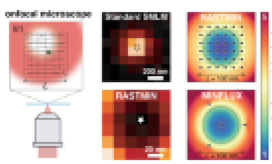
Sebastian Strauss, Philipp C. Nickels, … Ralf Jungmann



MINSTED nanoscopy enters the Ångström localization range

07 November 2022

Michael Weber, Henrik von der Emde, … Stefan W. Hell



An alternative to MINFLUX that enables nanometer resolution in a confocal microscope

30 June 2022

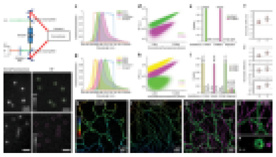
Luciano A. Masullo, Alan M. Szalai, … Fernando D. Stefani



Nanoparticles for super-resolution microscopy and single-molecule tracking

28 May 2018

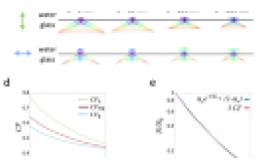
Dayong Jin, Peng Xi, … Antoine M. van Oijen



[**Nanoscale subcellular architecture revealed by multicolor three-dimensional salvaged fluorescence imaging**]

06 January 2020

Yongdeng Zhang, Lena K. Schroeder, … Joerg Bewersdorf



[**Three-dimensional total-internal reflection fluorescence nanoscopy with nanometric axial resolution by photometric localization of single molecules**]

22 January 2021

Alan M. Szalai, Bruno Siarry, … Fernando D. Stefani

- Article
- Open Access
- Published: 24 May 2023

Ångström-resolution fluorescence microscopy

- Susanne C. M. Reinhardt,
- Luciano A. Masullo,

- Isabelle Baudrexel,
- Philipp R. Steen,
- Rafal Kowalewski,
- Alexandra S. Eklund,
- Sebastian Strauss,
- Eduard M. Unterauer,
- Thomas Schlichthaerle,
- Maximilian T. Strauss,
- Christian Klein &
- …
- Ralf Jungmann

Show authors

Nature **volume 617**, pages 711–716 (2023)

- 30k Accesses
- 449 Altmetric
- Metrics details

Abstract

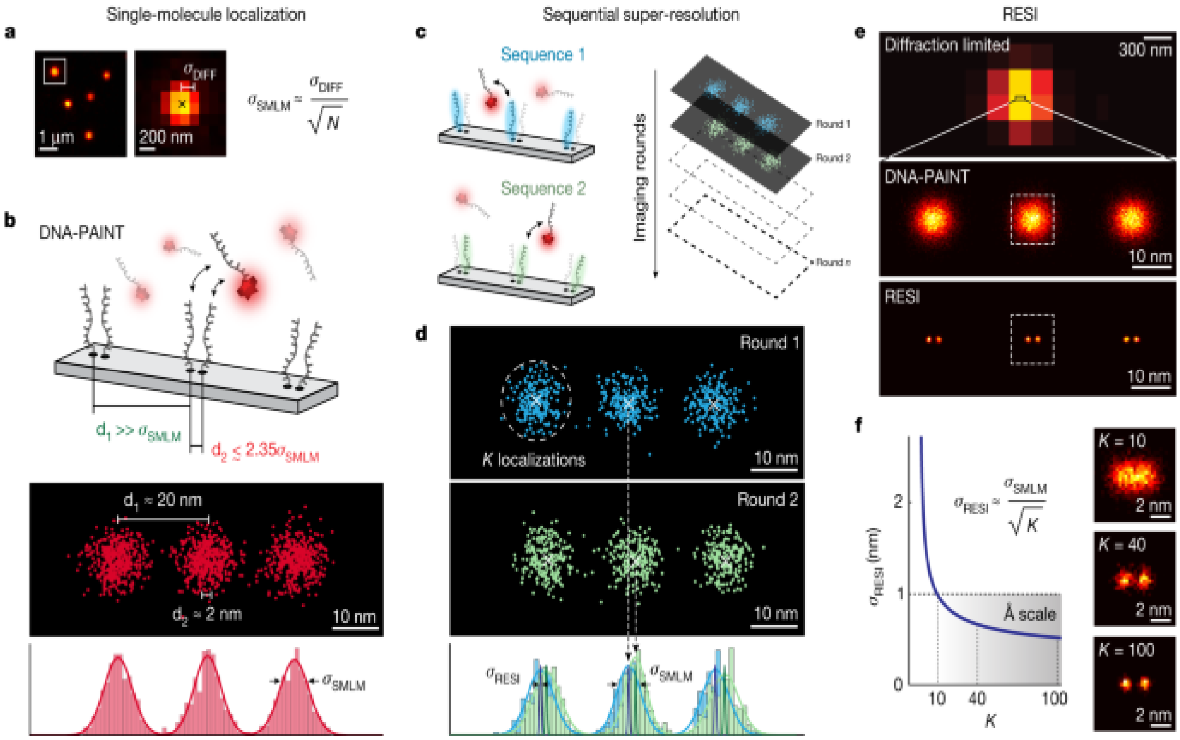
Fluorescence microscopy, with its molecular specificity, is one of the major characterization methods used in the life sciences to understand complex biological systems. Super-resolution approaches[1,2,3,4,5,6] can achieve resolution in cells in the range of 15 to 20 nm, but interactions between individual biomolecules occur at length scales below 10 nm and characterization of intramolecular structure requires Ångström resolution. State-of-the-art super-resolution implementations[7,8,9,10,11,12,13,14] have demonstrated spatial resolutions down to 5 nm and localization precisions of 1 nm under certain in vitro conditions. However, such resolutions do not directly translate to experiments in cells, and Ångström resolution has not been demonstrated to date. Here we introdue a DNA-barcoding method, resolution enhancement by sequential imaging (RESI), that improves the resolution of fluorescence microscopy down to the Ångström scale using

off-the-shelf fluorescence microscopy hardware and reagents. By sequentially imaging sparse target subsets at moderate spatial resolutions of >15 nm, we demonstrate that single-protein resolution can be achieved for biomolecules in whole intact cells. Furthermore, we experimentally resolve the DNA backbone distance of single bases in DNA origami with Ångström resolution. We use our method in a proof-of-principle demonstration to map the molecular arrangement of the immunotherapy target CD20 in situ in untreated and drug-treated cells, which opens possibilities for assessing the molecular mechanisms of targeted immunotherapy. These observations demonstrate that, by enabling intramolecular imaging under ambient conditions in whole intact cells, RESI closes the gap between super-resolution microscopy and structural biology studies and thus delivers information key to understanding complex biological systems.

Main

The localization precision ($\sigma_{\rm SMLM}$) of a target molecule in widefield single-molecule localization microscopy (SMLM)[15] is ultimately and fundamentally limited by the number of photons (N) collected per blinking event: ${\sigma }_{{\rm{SMLM}}}\approx \frac{{\sigma }_{{\rm{DIFF}}}}{\sqrt{N}}$ ($\sigma_{\rm DIFF}$ is the s.d. of the point spread function (PSF) of the optical imaging system[16]; Fig. 1a). Multiple localizations of the same target (Fig. 1b, top) are distributed around the true position due to their finite precision. Two or more points not resolvable by SMLM produce overlapping distributions of localizations, thus precluding unique assignment of localizations to respective targets (Fig. 1b, bottom). However, if each localization could be assigned to a specific target by colour, barcode or any other molecular identity, they could be unambiguously grouped per target[2].

Fig. 1: RESI concept.



a, In SMLM, σ_{SMLM} of a single dye scales with $\frac{{\sigma}_{{\rm{DIFF}}}}{\sqrt{N}}$, ultimately limiting the achievable spatial resolution. **b**, SMLM approaches such as DNA-PAINT feature approximately 10 nm spatial resolution (resolution approximated as full-width at half-maximum ≈ 2.35 σ_{SMLM}). Whereas targets separated by 20 nm (d_1) can thus be routinely resolved, objects spaced 2 nm apart (d_2) are unresolvable because the resulting distributions of localizations overlap. **c**, Using orthogonal DNA sequences (blue and green) and sequential acquisition as in Exchange-PAINT, localizations from targets spaced more closely than the SMLM resolution limit can be unambiguously assigned for each target. **d**, Combining all localizations per target (K) for each imaging round improves localization precision from s.d. (σ_{SMLM}) to s.e.m. (σ_{RESI}). **e**, As super-resolution revolutionized fluorescence microscopy, RESI results in another paradigm shift by reapplying the concept of localization to super-resolution data. **f**, Localization precision in RESI scales with $\frac{1}{\sqrt{K}}$, and thus resolution improvement in RESI is independent of σ_{SMLM}, reaching localization precision on the Ångström scale.

The centre of each group of localizations can be calculated with a precision far better than σ_{SMLM}. In essence, applying the principle of localization

microscopy to distinguishable groups of *K* super-resolved localizations, precision is increased from the s.d. (σ_{SMLM}) to the s.e.m. ($\frac{{\sigma }_{{\rm{SMLM}}}}{\sqrt{K}}$). Collecting an arbitrarily large number of localizations yields an arbitrary increase in precision. Notably, this increase in precision occurs regardless of the precision achieved in individual localizations (σ_{SMLM}).

We introduce a straightforward implementation of this concept using Exchange-PAINT[17], a variant of DNA-PAINT[18], for identical target molecules (Fig. 1c). DNA-PAINT uses the programmable, repetitive but transient binding of dye-labelled 'imager' strands to their complementary 'docking' strands on target molecules of interest[9,18]. The transient nature of the binding leads to an apparent 'blinking' of the target, necessary to perform SMLM. Exchange-PAINT uses orthogonal DNA barcodes combined with imaging and washing cycles to allow for sequential target multiplexing. In our implementation we 'multiplex' a single target species by separating it into multiple, sparser subsets. By imaging them sequentially, sufficiently spaced and isolated groups of localizations are measured. Determining the centre of each group of localizations yields a resolution enhancement (Fig. 1d). We call this implementation resolution enhancement by sequential imaging (RESI), and the resulting localizations RESI localizations.

By application of RESI in silico (Methods), we demonstrated a resolution improvement (Extended Data Fig. 1) over super-resolution akin to the improvement of super-resolved over diffraction-limited measurements (Fig. 1e). For routinely obtainable DNA-PAINT localization precision (approximately 3 nm) and number of localizations per target (in the order of hundreds), RESI could achieve precision well below one nanometre, thus entering the Ångström scale (Fig. 1f) according to ${\sigma }_{{\rm{RESI}}}=\frac{{\sigma }_{{\rm{SMLM}}}}{\sqrt{K}}$.

For an experimental proof of principle of RESI we used self-assembled DNA origami structures to precisely position orthogonal DNA strands[9,19]. We first designed DNA origami featuring two docking strands spaced 5 nm apart, a distance previously resolved with DNA-PAINT[7,9], to verify the accuracy and precision of RESI. Using two sequential imaging rounds and

an alignment procedure (Methods) to conduct RESI, we were able to accurately recapitulate the 5 nm point-to-point distance with precision improved by a factor of \(\sqrt{{K}_{{\rm{average}}}}=\sqrt{381}\approx 20\) (Extended Data Figs. 2 and 3).

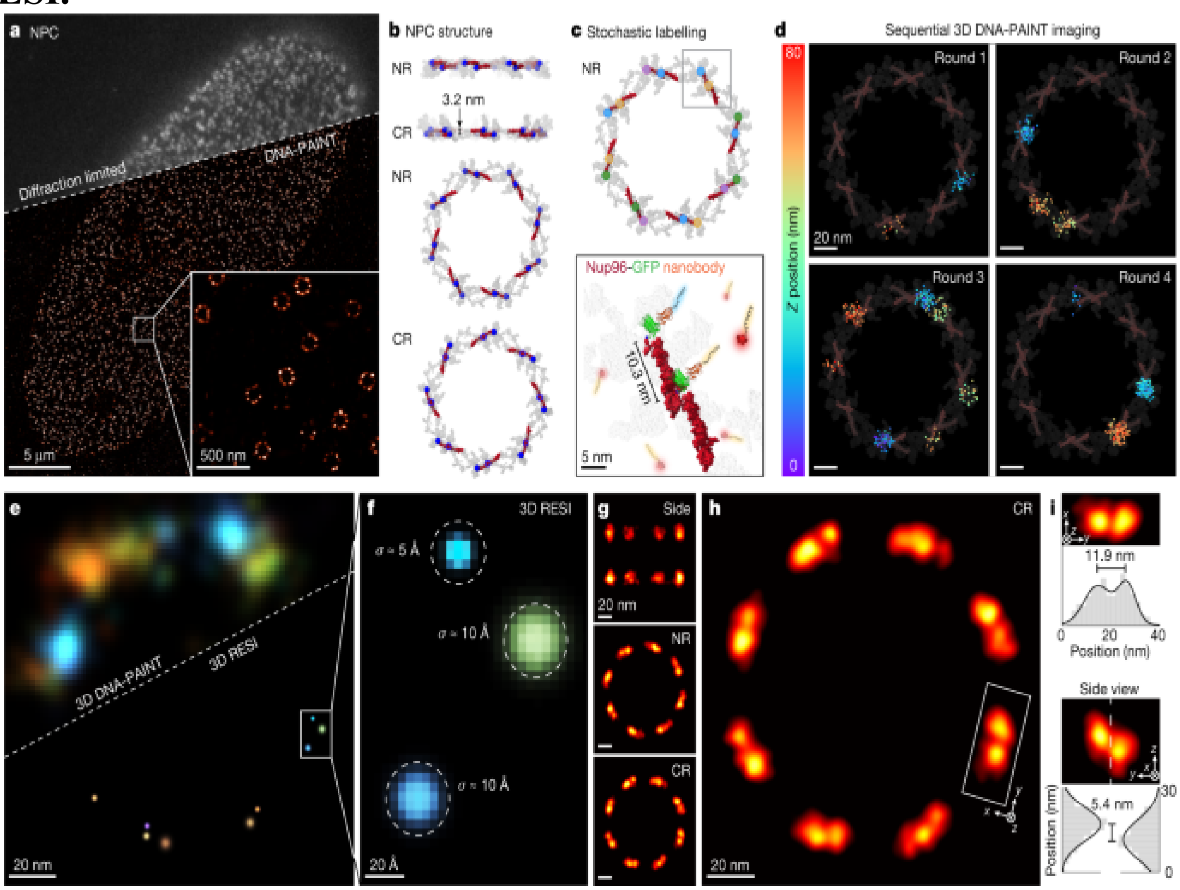
We next performed RESI in three dimensions (3D) using recently developed 3D DNA origami disk structures[20] and measured distances between docking strands of 2.5 ± 0.4 nm in *xy* and 11.3 ± 0.8 nm in *z*. This demonstrates that RESI resolution enhancement applies in all three dimensions, surpassing current state-of-the-art 3D super-resolution capabilities (Extended Data Figs. 4 and 5; for imaging parameters see Extended Data Table 1).

RESI resolves single nuclear pore complex proteins

To demonstrate the applicability of RESI in a cellular context, we next imaged structural proteins of the nuclear pore complex (NPC). As the major gatekeeper of nucleocytoplasmic transport, the NPC is a key target for structural biology research[21]. We furthermore chose the NPC as a model system because it has been well studied using a variety of imaging approaches, including cryo-electron microscopy (cryo-EM)[22], fluorescence microscopy and super-resolution techniques[23,24]. Figure 2a presents a typical diffraction-limited and DNA-PAINT image of Nup96 molecules (tagged with monomeric enhanced green fluorescent protein (mEGFP)) labelled with DNA-conjugated anti-GFP nanobodies. Nup96 is a structural NPC protein (part of the so-called Y-complex) present in eight pairs exhibiting an eight-fold symmetry on both cytoplasmic and nuclear rings, totalling 32 copies per NPC (Fig. 2b). Individual pairs of Nup96 proteins, spaced approximately 10 nm laterally and 3 nm axially, cannot be routinely resolved with current state-of-the-art super-resolution implementations[25,26,27,28]. To enable RESI, neighbouring copies of Nup96 proteins must be labelled with orthogonal DNA sequences. To this end, we opted for a stochastic labelling approach by incubating the sample with anti-GFP nanobodies, each conjugated with one out of four orthogonal sequences (Fig. 2c). We note that, with an increasing number of expected targets below the classical DNA-PAINT resolution limit, a larger number of orthogonal labelling sequences[29]—and thus

imaging rounds—is necessary to guarantee sufficiently spaced groups of localizations (for details on this requirement see Methods). Sequential 3D image acquisition in four rounds led to sufficiently spaced localization groups representing single Nup96 target molecules (Fig. 2d). Subsequent RESI super-localization of these groups allowed us to routinely visualize individual copies of Nup96 proteins (Fig. 2e). We note that this was achieved across the whole field of view (roughly 67 × 67 μm^2) totalling over 1,000 NPCs during a total image acquisition time of 100 min (see Extended Data Fig. 6 for representative data). The reconstructed RESI image features an average lateral localization precision of approximately 1 nm, representing a sixfold improvement over the individual DNA-PAINT acquisition rounds. We therefore achieved label-size-limited resolution, allowing us to resolve individual Nup96 molecules (Fig. 2f). Generally, label size not only limits spatial resolution but furthermore could lead to inaccuracies such as biased observed distances due to linkage errors.

Fig. 2: NPC proteins in whole cells resolved with Ångström precision by RESI.



a, Diffraction-limited and DNA-PAINT overview image of Nup96-mEGFP labelled with DNA-conjugated anti-GFP nanobodies. Zoomed-in view (bottom right) shows high labelling efficiency and image quality for standard DNA-PAINT conditions, recapitulating the eight-fold symmetry of the NPC. **b**, Cryo-EM structure representation of the location of Nup96 proteins (red; C-terminal mEGFP position marked in blue) as part of the Y-complex in nuclear and cytoplasmic rings (NR and CR, respectively). Adapted from PDB 7PEQ. Nup96 is present in 32 copies per NPC. **c**, To enable RESI, Nup96-mEGFP proteins were stochastically labelled with orthogonal DNA sequences by incubation of the sample with anti-GFP nanobodies, each conjugated with one of four orthogonal sequences (represented by blue, yellow, magenta and green dots). **d**, Sequential 3D imaging (colour represents *z* position) of the four labels yielded sufficiently spaced localization distributions. The average number of localizations per target is $K_{\text{average}} = 38$ (background represents cryo-EM structure from **b** for context). **e**, Comparison of 3D DNA-PAINT (top left) and 3D RESI (bottom right) for the same NPC illustrating improvement in spatial resolution by RESI. Localizations are rendered as gaussians with $\sigma_{\text{DNA-PAINT}}$ and σ_{RESI}, respectively. **f**, Localization precision (σ_{RESI}) as good as 5 Å was achieved by combining *K* localizations for each target, unambiguously resolving single Nup96 proteins. **g**, The 3D NPC cryo-EM structure was recapitulated using optical microscopy by applying a model-free average[30] of 1,217 NPCs from a single nucleus. **h**, RESI resolved adjacent Nup96 in a structural average by optical microscopy. **i**, Consistent with the cryo-EM structure (taking into account linkage error arising from label size), adjacent Nup96 proteins were spaced 11.9 ± 1.2 nm apart laterally (top) and 5.4 ± 0.4 nm axially (bottom).

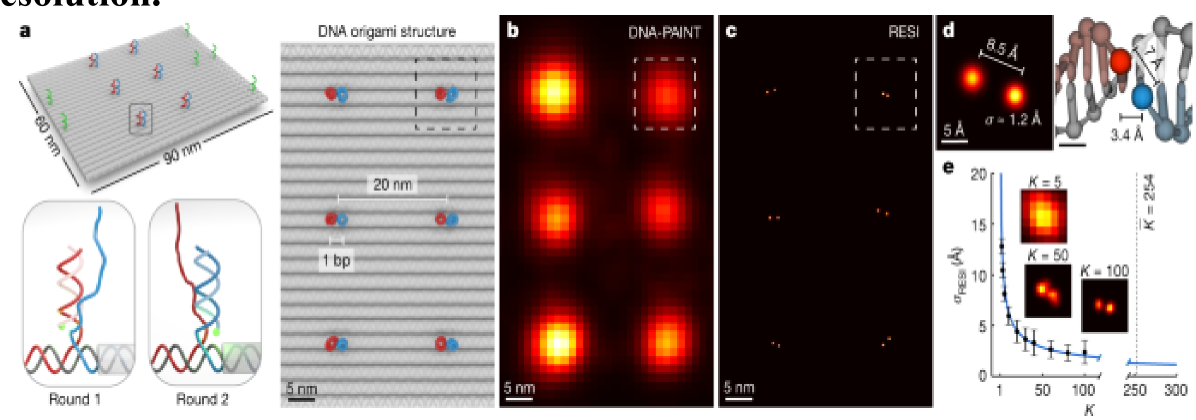
We then performed unbiased 3D averaging of 1,217 NPCs using a recently developed model-free approach for SMLM data[30]. The resulting 3D average (Fig. 2g) not only allows recapitulation of the eight-fold symmetry of Nup96 in both cytoplasmic and nuclear rings (which has previously been achieved with super-resolution[23,24,25,26,27,28]), but enables resolution of individual Nup96 proteins in a structural average (Fig. 2h). Enabled by RESI's unprecedented spatial resolution, we were able to recapitulate distances of Nup96 proteins of 11.9 ± 1.2 nm laterally and 5.4 ± 0.4 nm axially from the structural average image (Fig. 2i). Both lateral and axial orientation, as well

as tilt, of Nup96 pairs are consistent with cryo-EM data[22]. We resolved this spatial arrangement for most Nup96 protein pairs (Extended Data Fig. 7), which was previously out of reach for optical microscopy.

Imaging DNA bases at Ångström resolution

To assay the ultimately achievable spatial resolution by RESI, we designed a flat, rectangular DNA origami structure featuring six pairs (spaced 20 nm apart) of directly adjacent orthogonal docking strands at a distance of only one DNA base pair (red and blue strands in Fig. 3a). This yielded a designed in-plane distance of around 7 Å along the backbone of one strand of the DNA double helix[31]. The structures also contain DNA-PAINT docking strands for precise alignment between sequential imaging rounds (green strands in Fig. 3a). State-of-the-art DNA-PAINT image acquisition[32] at approximately 5 nm spatial resolution yielded six localization clouds in a 20 nm grid arrangement but failed to resolve the individual docking strands at subnanometre single-base-pair distances (Fig. 3b).

Fig. 3: RESI resolves the distance of single DNA base pairs at Ångström resolution.



a, DNA origami with docking strands spaced by a single base pair (bp; red and blue strands, with alignment markers in green) provided a platform to demonstrate the highest resolution achievable by RESI. **b**, DNA-PAINT resolved 20 nm spacing but the resolution was insufficient to distinguish individual docking sites, spaced one base apart. **c**, RESI resolves the adjacent docking strands. **d**, A Euclidean distance of 8.5 ± 1.7 Å was calculated from individual localizations with an average precision of 1.2 Å

(left) for the single-base-pair backbone distance, which is within 1 s.d. of the expected value of roughly 7 Å (right). **e**, Experimental localization precision in RESI is in good agreement with $\frac{{\sigma }_{{\rm{SMLM}}}}{\sqrt{K}}$ (blue line, K), yielding an average localization precision of 1.3 Å for the experimental data from all $n = 42$ DNA origami (insets correspond to exemplary point pair in **d**). Error bars represent mean ± 1 s.d.

Remarkably, RESI resolves the individual docking strand positions (Fig. 3c) in all DNA origami structures. We note that RESI achieved this in an image acquisition time of 100 min featuring an approximately 67 × 67 μm^2 field of view containing more than 2,000 DNA origami structures (see Extended Data Fig. 8 for representative DNA origami structures). RESI allows us to routinely resolve strands spaced apart by only one DNA base pair. Strikingly, we measured a distance of 8.5 ± 1.7 Å between two single docking strands in an individual DNA origami structure (Fig. 3d). This demonstrates an unprecedented resolution in optical microscopy by distinguishing structures closer than one nanometre. We note that our resolution claim is based on the most fundamental and strict definition: the ability to spatially distinguish point objects. We measured a distance of 9.5 ± 2.6 Å between adjacent docking strands in an average of 42 DNA origami (Extended Data Fig. 9), which is within 1 s.d. of the expected backbone distance[31] of around 7 Å.

To quantify resolution gain, we calculated RESI localizations for different values of K underlying DNA-PAINT localizations (Methods). We demonstrate that the effective localization precision scales as ${\sigma }_{{\rm{RESI}}}=\frac{{\sigma }_{{\rm{SMLM}}}}{\sqrt{K}}$, yielding an average localization precision of 1.3 Å for an average $K = 254$ (Fig. 3e), experimentally confirming the in silico results (Fig. 1f). RESI not only yields virtually unlimited numbers of localizations per target, but also avoids detrimental photophysical effects caused by spatial proximity of fixed-dye labels because, in DNA-PAINT imaging, two adjacent dyes are never present simultaneously. It has recently been reported[33] that, at sub-10-nm distances, photophysical near-field interactions play a major role in modulation of photoswitching kinetics, thus effectively preventing fixed-dye SMLM techniques from accessing this resolution scale. This ultimately limits the achievable resolution of even the most photon-efficient techniques

available for single-molecule localization, such as MINFLUX or MINSTED, despite their subnanometre precision, unless combined with DNA-PAINT. The experimentally demonstrated subnanometre resolution illustrates the capacity of RESI to enable structural biology studies using DNA-based imaging at hitherto elusive scales.

CD20 receptor organization

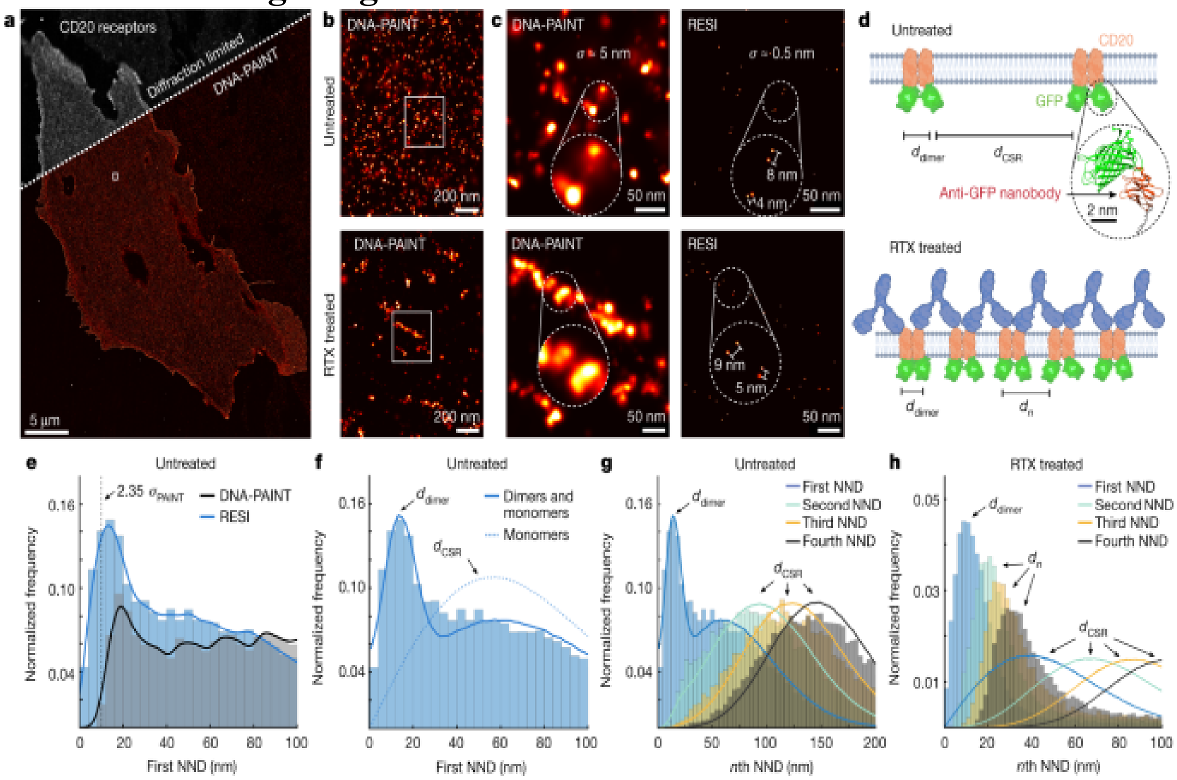
Finally, we applied RESI to address and resolve a cell-biological question currently under debate that has so far been beyond reach for both cryo-EM in a native cellular context and incumbent super-resolution techniques. Specifically we studied the organization of CD20 membrane receptors, which are prime targets for therapeutic antibody treatment of B cell-derived blood cancers and autoimmune diseases[34].

In the case of the most frequently used therapeutic anti-CD20 antibody, rituximab (RTX), the spatial rearrangement of CD20 in the cell membrane is thought to play an important role in its efficacy[35,36]. Recent cryo-EM studies detected CD20 as a dimer in complex with two individual RTX-fragment antigen-binding regions[37,38], suggesting a linear chain-like assembly of CD20 in the presence of the full antibody[38]. On the other hand, when incubated with the full RTX antibody, a trimeric ring of alternating RTX molecules and CD20 dimers was detected in EM images[37]. The fact that cryo-EM experiments are performed in detergent solution raises the question about which molecular arrangements are actually present in the cell. Currently CD20 organization when bound to full RTX antibodies in intact cells cannot be assessed, thus precluding the investigation of whether CD20 clusters are of linear or circular nature. Moreover, even though in vitro studies showed that CD20 dimers can form without antibody binding, the quantitative assessment of CD20 dimerization in untreated cells is currently limited.

Here we applied RESI to study the molecular arrangement of CD20 in Chinese hamster ovary (CHO) cells transiently transfected with mEGFP-CD20, using four rounds of probe exchange in a total imaging time of 4.4 h. In the diffraction-limited overview and DNA-PAINT super-resolution image of untreated cells, CD20 appeared homogeneously distributed (Fig. 4a,b

(top) and Extended Data Fig. 10a) whereas RTX-treated cells exhibited apparent CD20 clusters (Fig. 4b (bottom) and Extended Data Figs. 11a and 12a).

Fig. 4: RESI shows CD20 receptor (re)organization at subnanometre precision following drug treatment.



a, Diffraction-limited and DNA-PAINT overview image of CHO cells expressing mEGFP-CD20 labelled with anti-GFP nanobodies. **b**, Zoomed-in DNA-PAINT image showing apparently randomly distributed CD20 receptors for untreated cells (top) and clustered receptor arrangement for RTX-treated cells (bottom). **c**, Comparison of DNA-PAINT and RESI for both untreated and RTX-treated cells showing sub-10-nm-spaced receptor pairs in the RESI images, which are unresolvable with DNA-PAINT. **d**, RESI data suggest that CD20 proteins occur in dimers (spaced at d_{dimer}), which are in turn distributed according to complete spatial randomness (CSR; distances between dimers, d_{CSR}) in untreated cells. Chains of dimers were observed following administration of RTX. **e**, Whole-cell analysis of first NNDs of CD20 receptors (histograms of distances and kernel density estimation are shown). Only RESI, but not DNA-PAINT, allows the routine

detection of sub-10-nm distances between proteins. Whereas DNA-PAINT overestimates dimer distance, RESI shows a label-limited distance of 13.5 nm (see main text for discussion). **f**, Fitting RESI NND data from **e** to a numerical model reveals CD20 dimers and monomers. **g**, CD20 receptors in untreated cells showed second to fourth NNDs consistent with CSR, thus excluding the presence of higher-order protein complexes. **h**, CD20 receptors in RTX-treated cells, however, showed first to fourth NNDs, inconsistent with complete spatial randomness.

Comparison of DNA-PAINT and RESI for both untreated and RTX-treated cells shows sub-10-nm-spaced CD20 pairs in the RESI images (Fig. 4c, right) that were unresolvable with DNA-PAINT (Fig. 4c, left). RESI images suggest that CD20 is present in dimers and chain-like, higher-order structures in untreated and RTX-treated cells, respectively (Fig. 4d).

To quantitatively assess the existence of dimers in untreated cells, we performed first nearest-neighbour distance (NND) analysis for both DNA-PAINT and RESI data, demonstrating nonrandom distributions in both cases (Fig. 4e). RESI at 1 nm localization precision shows a substantial fraction of sub-10-nm distances in the NND histogram, which enables quantitative assessment of the degree of CD20 dimerization. We performed numerical simulations and a least-squares fit (Methods) that yielded a composition of 53 ± 1% monomers and 47 ± 1% dimers with average intradimer distance of 13.5 ± 0.3 nm (Fig. 4f, solid line). For comparison, we simulated NND distributions corresponding to a population of 100% monomers (Fig. 4f, dotted line), further demonstrating that CD20 molecules are not present solely as monomers. Because all NND distributions except for the first order are consistent with a complete spatial random (CSR) distribution at the experimentally measured density, we exclude the presence of higher-order assemblies for untreated CD20 (Fig. 4g). Our findings present quantitative experimental evidence that CD20 exists as dimers in an intact cell membrane.

By contrast, RESI analysis of CD20 in RTX-treated cells yielded first to fourth NND distributions inconsistent with a CSR model (Fig. 4h and Extended Data Fig. 12d,e). This suggests a higher-order arrangement of CD20 molecules after RTX treatment and confirms recent cryo-EM-derived models[37,38].

Finally we probed the existence of hexameric, ring-like arrangements by comparison with numerical simulations (Extended Data Fig. 13). The characteristics of the experimentally detected CD20 clusters suggest the absence of isolated hexamers and support the hypothesis of predominantly linear, chain-like structures (Extended Data Fig. 13h).

Discussion

We introduce RESI, a conceptually new approach in SMLM to improve the spatial resolution of optical microscopy to the Ångström scale. RESI achieves this by combining multiple localizations from single targets to obtain a ‘super-super-resolution’ image after separating their localizations by sequential imaging (for example, using DNA-barcoded probes).

In this way RESI precision—and thus resolution—scales not only with the number of photons (*N*) detected per localization but also with the number of localizations (*K*) acquired per target. RESI thus provides a new precision scaling law: ${\sigma }_{{\rm{RESI}}}=\frac{{\sigma }_{{\rm{SMLM}}}}{\sqrt{K}}\approx \frac{{\sigma }_{{\rm{DIFF}}}}{\sqrt{K}\sqrt{N}}$. This applies if a sufficiently large number of orthogonal labelling sequences and thus imaging rounds guarantee adequately spaced groups of localizations (Extended Data Fig. 14). Importantly, resolution enhancement is isotropic in three dimensions. For our current experimental implementation, RESI approaches structural biology resolution with an all-optical approach in intact cells using off-the-shelf labelling reagents and a simple inverted fluorescence microscope operated under ambient conditions. We were able to experimentally demonstrate Ångström spatial resolution below the physical size of a dye. This was achieved due to three specific advantages of DNA-PAINT leading to unbiased target sampling: (1) the rotational flexibility of the target-bound docking strand (even in the case of longer repetitive-sequence motifs[32]); (2) the freely rotating dipole of the dye attached to the imager sequence; and (3) the fact that two adjacent imagers are never present simultaneously.

Furthermore, because RESI images are not obtained from single localizations but from groups of localizations per target, the method presents a uniquely robust feature compared with other SMLM techniques: it shifts

the focus from enhancement of only optical precision ($\sigma_{\rm OPT}$) to improvement in overall precision ($\sigma_{\rm SMLM}\approx \sqrt{\sigma_{\rm OPT}^{2}+\sigma_{\rm MEC}^{2}}$) by averaging out the uncertainty effects of mechanical instability ($\sigma_{\rm MEC}$), provided the latter is normally distributed.

With RESI we measured areas of 67 × 67 μm^2 in 100 min, making it applicable as a sufficiently high-throughput tool for cell biology. Resolving receptor patterns at single-protein resolution could enable ‘spatial diagnostics’ as a prescreening method for personalized treatments, and serve as a tool for biomedical discovery of patterned therapeutics—for example, by guiding drug design principles.

RESI performance and accuracy could be further improved by advances in intramolecular labelling approaches such as orthogonal unnatural amino acids[39]. RESI is thus poised to close the gap between 3D fluorescence super-resolution microscopy in whole intact cells and cryo-EM structural studies of individual supramolecular complexes, introducing a paradigm shift in fluorescence imaging by pushing optical microscopy to Ångström resolutions.

Methods

Materials

Unmodified DNA oligonucleotides, as well as DNA oligonucleotides modified with C3-azide and Cy3B, were purchased from MWG Eurofins and Metabion. The M13mp18 and p7560 scaffold was obtained from Tilibit. Magnesium chloride (1 M, no. AM9530G), sodium chloride (5 M, no. AM9759), ultrapure water (no. 10977-035), Tris (1 M, pH 8.0, no. AM9855G), EDTA (0.5 M, pH 8.0, no. AM9260G) and 10× PBS (no. 70011051) were purchased from Thermo Fisher Scientific. BSA (no. A4503-10G) was ordered from Sigma-Aldrich. Triton X-100 (no. 6683.1) was purchased from Carl Roth. Sodium hydroxide (no. 31627.290) was purchased from VWR. Paraformaldehyde (no. 15710) and glutaraldehyde (no. 16220) were obtained from Electron Microscopy Sciences. Tween-20 (no. P9416-50ML), glycerol (no. 65516-500 ml), methanol (no. 32213-

2.5L), protocatechuate 3,4-dioxygenase pseudomonas (PCD, no. P8279), 3,4-dihydroxybenzoic acid (PCA, no. 37580-25G-F) and (±)-6-hydroxy-2,5,7,8-tetra-methylchromane-2-carboxylic acid (trolox, no. 238813-5G) were ordered from Sigma-Aldrich. Neutravidin (no. 31000) was purchased from Thermo Fisher Scientific. Biotin-labelled BSA (no. A8549) and sodium azide (no.769320) were obtained from Sigma-Aldrich. Coverslips (no. 0107032) and glass slides (no. 10756991) were purchased from Marienfeld and Thermo Fisher Scientific, respectively. Fetal bovine serum (FBS, no. 10500-064), 1× PBS (pH 7.2, no. 20012-019), 0.05% trypsin-EDTA (no. 25300-054), salmon sperm DNA (no. 15632011), OptiMEM (no. 31985062) and Lipofectamine LTX (no. A12621) were purchased from Thermo Fisher Scientific. Gold nanoparticles (90 nm, no. G-90-100) were ordered from Cytodiagnostics. Nanobodies against GFP (clone 1H1) with a single ectopic cysteine at the C terminus for site-specific conjugation were purchased from Nanotag Biotechnologies. DBCO-PEG4-Maleimide (no. CLK-A108P) was purchased from Jena Bioscience.

Buffers

The following buffers were used for sample preparation and imaging.

- Buffer A: 10 mM Tris pH 8.0, 100 mM NaCl and 0.05% Tween-20
- Buffer B: 10 mM $MgCl_2$, 5 mM Tris-HCl pH 8.0, 1 mM EDTA and 0.05% Tween-20 pH 8.0
- Buffer C: 1× PBS, 1 mM EDTA, 500 mM NaCl pH 7.4, 0.02% Tween, optionally supplemented with 1× trolox, 1× PCA and 1× PCD
- Blocking buffer: 1× PBS, 1 mM EDTA, 0.02% Tween-20, 0.05% NaN_3, 2% BSA, 0.05 mg ml^{-1} sheared salmon sperm DNA
- Two-dimensional (2D) DNA origami folding buffer: 10 mM Tris, 1 mM EDTA, 12.5 mM $MgCl_2$ pH 8.0
- 3D DNA origami folding buffer: 5 mM Tris, 1 mM EDTA, 5 mM NaCl, 20 mM MgCl2 pH 8.0

- 1× TA buffer: 40 mM Tris pH 8.0, 20 mM acetic acid

PCA, PCD and trolox

Trolox (100×) was made by the addition of 100 mg of trolox to 430 µl of 100% methanol and 345 µl of 1 M NaOH in 3.2 ml of water. PCA (40×) was made by mixing 154 mg of PCA in 10 ml of water and NaOH and adjustment of pH to 9.0. PCD (100×) was made by the addition of 9.3 mg of PCD to 13.3 ml of buffer (100 mM Tris-HCl pH 8.0, 50 mM KCl, 1 mM EDTA, 50% glycerol).

DNA-PAINT docking and imager sequences

Four orthogonal DNA sequence motifs were used to label targets in four RESI rounds. The docking strands were 5xR1 (TCCTCCTCCTCCTCCTCCT), 5xR2 (ACCACCACCACCACCACCA), 7xR3 (CTCTCTCTCTCTCTCTCTC) and 7xR4 (ACACACACACACACACACA). The respective imagers were R1 (AGGAGGA-Cy3B), R2 (TGGTGGT-Cy3B), R3 (GAGAGAG-Cy3B) and R4 (TGTGTGT-Cy3B). The design of 2D RESI origami required extension of the R1 site at the 5′ end such that the adjacent R1 and R3 docking strands could be spaced apart by a single base pair. Thus, the docking strand 5′ 5xR1 (TCCTCCTCCTCCTCCTCCT) and the 5′ R1 imager (Cy3B-AGGAGGA) were used rather than the 3′ versions for both 2D DNA origamis.

DNA origami self-assembly (2D)

All 2D DNA origami structures were designed in caDNAno[40]. Self-assembly of DNA origami was accomplished in a one-pot reaction mix with a total volume of 40 µl, consisting of 10 nM scaffold strand (for sequence, see Supplementary Data 2), 100 nM folding staples (Supplementary Data 1), 500 nM biotinylated staples (Supplementary Data 1) and 1 µM staple strands with docking site extensions (Supplementary Data 1) in 2D DNA origami folding buffer. The reaction mix was then subjected to a thermal annealing ramp using a thermocycler. First, it was incubated at 80 °C for 5 min, cooled using a temperature gradient from 60 to 4 °C in steps of 1 °C per 3.21 min and finally held at 4 °C.

DNA origami self-assembly (3D)

The 3D DNA origami disk structure was designed in caDNAno[40]. Self-assembly of the DNA origami disk was accomplished in a one-pot reaction mix of 50 µl total volume, consisting of 20 nM scaffold strand p7560 (for sequence, see Supplementary Data 3), 200 nM core folding staples (Supplementary Data 1), 200 nM staple sequences without handle extension (Supplementary Data 1), 500 nM biotinylated staples (Supplementary Data 1), 2 µM staple strands with R4 docking site extensions and 4 µM staple strands with R1 or R3 docking site extensions (Supplementary Data 1) in 3D DNA origami folding buffer. The reaction mix was then subjected to a thermal annealing ramp using a thermocycler. It was first incubated at 80 °C for 5 min then cooled using a temperature gradient from 60 °C to 20 °C in steps of 1 °C h^{-1} and finally held at 20 °C.

DNA origami purification

After self-assembly, structures were purified by agarose gel electrophoresis (1.5% agarose, 1× TA, 10 mM $MgCl_2$, 0.5× SybrSafe) at 4.5 V cm^{-1} for 1.5 h. Gel bands were cut, crushed and the origami stored in low-binding Eppendorf tubes at −20 °C.

DNA origami sample preparation and imaging

For sample preparation, a bottomless six-channel slide (ibidi, no. 80608) was attached to a coverslip. First, 80 µl of biotin-labelled BSA (1 mg ml^{-1}, dissolved in buffer A) was flushed into the chamber and incubated for 5 min. The chamber was then washed with 360 µl of buffer A. A volume of 100 µl of neutravidin (0.1 mg ml^{-1}, dissolved in buffer A) was then flushed into the chamber and allowed to bind for 5 min. After washing with 180 µl of buffer A and subsequently with 360 µl of buffer B, 80 µl of biotin-labelled DNA structures (approximately 200 pM) in buffer B was flushed into the chamber and incubated for 5 min. For measurement of the DNA origami disk, additional 2D DNA origami structures with 12 target sites[9] spaced 20 nm apart were incubated together, with the 3D disk origami serving as fiducials for drift correction. After DNA origami incubation the chamber was washed

with 540 µl of buffer B. For DNA origami disk structures, 150 µl of gold nanoparticles (diluted 1:10 in buffer B) was flushed through and incubated for 5 min before washing with 540 µl of buffer B. Finally, 180 µl of the imager solution in buffer B was flushed into the chamber. The chamber remained filled with imager solution and imaging was then performed. Between imaging rounds, the sample was washed three times with 1 ml of buffer B until no residual signal from the previous imager solution was detected. Then, the next imager solution was introduced. For RESI, two imaging rounds were performed with imagers R1 and R4 present in round 1 and the imagers R3 and R4 in round 2 (R1 and R3 probe the sites of interest for RESI and R4 serves alignment purposes).

Nanobody–DNA conjugation

Nanobodies were conjugated as described previously[32]. Unconjugated nanobodies were thawed on ice, then 20-fold molar excess of bifunctional DBCO-PEG4-Maleimide linker was added and reacted for 2 h on ice. Unreacted linker was removed by buffer exchange to PBS using Amicon centrifugal filters (10,000 MWCO). The DBCO-modified nanobodies were reacted with 5× molar excess of azide-functionalized DNA (R1, R2, R3 and R4) overnight at 4 °C. Unconjugated protein and free DNA were removed by anion exchange chromatography using an ÄKTA pure system equipped with a Resource Q 1 ml column.

Cell culture

CHO cells (CCL-61, ATCC) were cultured in Gibco Ham's F-12K (Kaighn's) medium supplemented with 10% FBS (no. 11573397, Gibco). U2OS-CRISPR-Nup96-mEGFP cells (a gift from the Ries and Ellenberg laboratories) were cultured in McCoy's 5A medium (Thermo Fisher Scientific, no. 16600082) supplemented with 10% FBS. Cells were passaged every 2–3 days using trypsin-EDTA.

Nup96 EGFP imaging

U2OS-CRISPR-Nup96-mEGFP cells were seeded on ibidi eight-well high glass-bottom chambers (no. 80807) at a density of 30,000 cm^{-2}. Cells were

fixed with 2.4% paraformaldehyde in PBS for 30 min at room temperature. After fixation, cells were washed three times with PBS. Gold nanoparticles (200 µl) were incubated for 5 min and washed three times with PBS. Blocking and permeabilization were performed with 0.25% Triton X-100 in blocking buffer for 90 min. After washing with PBS, cells were incubated with 100 nM anti-GFP nanobodies in blocking buffer for 60 min at room temperature. To enable RESI, the nanobody solution consisted of 25 nM R1, R2, R3 and R4 docking-strand-coupled anti-GFP nanobodies with a total nanobody concentration of 100 nM. Unbound nanobodies were removed by washing three times with PBS, followed by washing once with buffer C for 10 min. Postfixation was performed with 2.4% paraformaldehyde in PBS for 15 min. After washing 3× with PBS, the imager solution in buffer C was flushed into the chamber. Between imaging rounds the sample was washed with 1–2 ml of PBS until no residual signal from the previous imager solution was detected. Then, the next imager solution was introduced. First, imagers R1, R2, R3 and R4 were added simultaneously to the sample to perform a standard DNA-PAINT measurement; then, RESI imaging was conducted via four subsequent imaging rounds with only one of the imagers.

Cloning

mEGFP-Alfa-CD20 was cloned by insertion of Alfa-CD20 into the mEGFP-C1 plasmid (no. 54759, Addgene). An Alfa-CD20 gblock (obtained from Integrated DNA Technologies) was amplified with primers cggcatggacgagct and gtacaagtccgga and, after cutting with restriction enzymes BsrGI and BamHI, Gibson assembly was performed (2× mix, NEB).

mEGFP-CD20 imaging

CHO cells were seeded on ibidi eight-well high glass-bottom chambers (no. 80807) the day before transfection at a density of 15,000 cm^{-2}. Transfection with mEGFP-CD20 was carried out with Lipofectamine LTX as specified by the manufacturer. CHO cells were allowed to express mEGFP-CD20 for 16–24 h. Then, the medium was replaced with fresh F-12K medium + 10% FBS (in the untreated case) or with F-12K medium + 10% FBS + 10 ug ml^{-1} RTX-Alexa 647 (a gift from Roche Glycart) (in the RTX-treated case), followed by incubation for 30 min. After washing two times with fresh

medium for 5 min, cells were fixed with 250 µl of prewarmed 4% PFA + 0.1% glutaraldehyde in PBS for 15 min. CHO cells were washed three times with PBS and quenched with 100 mM Tris pH 8.0 for 5 min. Permeabilization was carried out for 5 min with 0.2% Triton X-100 in PBS, followed by three washes with PBS. Cells were blocked in blocking buffer for 1 h at room temperature (RT). Anti-GFP nanobodies were incubated at a total concentration of 25 nM overnight at 4 °C; for RESI with four rounds this yielded 6.25 nM each of GFP-Nb-R1/2/3/4. After washing three times with PBS at RT for 15 min, cells were postfixed with 4% PFA at RT for 10 min followed by washing and postfixation as described above. Gold nanoparticles (90 nm) were diluted 1:3 in PBS and incubated for 10 min at RT and the sample was washed two times with PBS to remove unbound gold. The imager solution in buffer C for the first round was incubated for 5 min and then replaced with fresh imager, after which the first acquisition round was started. Between imaging rounds the sample was washed with at least 2 ml of PBS until no residual signal from the previous imager solution was detected. Then, the next imager solution was introduced. RESI imaging was conducted via four subsequent imaging rounds with only one of the imagers. In the final imaging round, imagers R1, R2, R3 and R4 were added simultaneously to the sample to perform a standard DNA-PAINT measurement.

Microscopy setup

Fluorescence imaging was carried out using an inverted microscope (Nikon Instruments, Eclipse Ti2) with the Perfect Focus System, applying an objective-type TIRF configuration equipped with an oil-immersion objective (Nikon Instruments, Apo SR TIRF ×100/numerical aperture 1.49, oil). A 560 nm laser (MPB Communications, 1 W) was used for excitation. The laser beam was passed through a cleanup filter (Chroma Technology, no. ZET561/10) and coupled to the microscope objective using a beam splitter (Chroma Technology, no. ZT561rdc). Fluorescence was spectrally filtered with an emission filter (Chroma Technology, nos. ET600/50m and ET575lp) and imaged on an sCMOS camera (Andor, Zyla 4.2 Plus) without further magnification, resulting in an effective pixel size of 130 nm (after 2 × 2 binning). The readout rate was set to 200 MHz. Images were acquired by choosing a region of interest of size 512 × 512 pixels. 3D imaging was

performed using a cylindrical lens (Nikon Instruments, N-STORM) in the detection path. Raw microscopy data were acquired using µManager[41] (v.2.0.1). Total internal reflection illumination was used for 2D and 3D DNA origami data, as well as for CD20 acquisition. Highly inclined and laminated optical sheet (HILO) illumination was employed for the acquisition of NPC data. Detailed imaging conditions for the respective experiments are shown in Extended Data Table 1.

Imaging parameters and duration

Due to target and sample heterogeneity the optimal imager concentration, c, used to achieve sparse blinking varies. Here we used concentrations from 100 pM (Nup96) to 800 pM (DNA origami). Optimal imager concentrations were determined visually for each sample. Concentrations were altered until blinking was frequent but sufficiently sparse to achieve good DNA-PAINT resolution.

The average number of expected binding events per binding site during a DNA-PAINT measurement is given by the duration of the measurement $t_{\rm measurement}$ and the mean dark time $\tau_{\rm dark}$ (defined as ${\tau}_{{\rm{dark}}}=\frac{1}{{k}_{{\rm{on}}}\times c}$), with $k_{\rm on}$ being the on-rate of a given imager strand) as:

$${n}_{{\rm{b}}{\rm{i}}{\rm{n}}{\rm{d}}{\rm{i}}{\rm{n}}{\rm{g}}{\rm{e}}{\rm{v}}{\rm{e}}{\rm{n}}{\rm{t}}{\rm{s}}}=\left(\frac{{t}_{{\rm{m}}{\rm{e}}{\rm{a}}{\rm{s}}{\rm{u}}{\rm{r}}{\rm{e}}{\rm{m}}{\rm{e}}{\rm{n}}{\rm{t}}}}{{\tau }_{{\rm{d}}{\rm{a}}{\rm{r}}{\rm{k}}}}\right)={t}_{{\rm{m}}{\rm{e}}{\rm{a}}{\rm{s}}{\rm{u}}{\rm{r}}{\rm{e}}{\rm{m}}{\rm{e}}{\rm{n}}{\rm{t}}}\times {k}_{{\rm{o}}{\rm{n}}}\times c.$$

The average number of localizations per binding event is given by the mean bright time $\tau_{\rm bright}$ and camera exposure time $t_{\rm exposure}$ as

$${n}_{{\rm{locs}}{\rm{per}}{\rm{binding}}{\rm{event}}}=\left(\frac{{\tau }_{{\rm{bright}}}}{{t}_{{\rm{exposure}}}}\right).$$

Therefore, the average number of localizations expected per binding site over the course of the measurement is

$${n}_{{\rm{loc}}}=\left(\frac{{t}_{{\rm{measurement}}}}{{\tau }_{{\rm{dark}}}}\right)\times \left(\frac{{\tau }_{{\rm{bright}}}}{{t}_{{\rm{exposure}}}}\right)={t}_{{\rm{measurement}}}\times {k}_{{\rm{on}}}\times c\times \left(\frac{{\tau }_{{\rm{bright}}}}{{t}_{{\rm{exposure}}}}\right).$$

It follows that the total acquisition time necessary to collect n_{loc} localizations is, on average,

$${t}_{{\rm{m}}{\rm{e}}{\rm{a}}{\rm{s}}{\rm{u}}{\rm{r}}{\rm{e}}{\rm{m}}{\rm{e}}{\rm{n}}{\rm{t}}}=\frac{{t}_{{\rm{e}}{\rm{x}}{\rm{p}}{\rm{o}}{\rm{s}}{\rm{u}}{\rm{r}}{\rm{e}}}\times {n}_{{\rm{l}}{\rm{o}}{\rm{c}}}}{{\tau }_{{\rm{b}}{\rm{r}}{\rm{i}}{\rm{g}}{\rm{h}}{\rm{t}}}{\times k}_{{\rm{o}}{\rm{n}}}\times c}.$$

The necessary number of localizations, n_{loc}, is calculated using $({\sigma }_{{\rm{RESI}}}=\frac{{\sigma }_{{\rm{DNA}} \mbox{-}{\rm{PAINT}}}}{\sqrt{{n}_{{\rm{loc}}}}}$), and thus $({n}_{{\rm{loc}}}={\left(\frac{{\sigma }_{{\rm{DNA}} \mbox{-}{\rm{PAINT}}}}{{\sigma }_{{\rm{RESI}}}}\right)}^{2}$) with the DNA-PAINT localization precision $\sigma_{DNA\text{-}PAINT}$.

For expected imager concentrations between 50 and 800 pM, exposure times between 100 and 200 ms and kinetics reported previously[32], the times required to collect 16 localizations (1 nm RESI precision given $\sigma_{DNA\text{-}PAINT}$ = 4 nm) vary between 42 s (R2, 800 pM, 100 ms exposure time) and 314 min (R5, 50 pM, 200 ms exposure time).

DNA-PAINT analysis

Raw fluorescence data were subjected to super-resolution reconstruction using the Picasso software package[9] (latest version available at https://github.com/jungmannlab/picasso). Drift correction was performed with a redundant cross-correlation and gold particles as fiducials for cellular

experiments, or with single DNA-PAINT docking sites as fiducials for origami experiments.

Channel alignment

Alignment of subsequent imaging rounds was performed iteratively in Picasso[9], starting with a redundant cross-correlation and followed by gold fiducial alignment for cellular experiments. Every DNA origami was equipped with additional DNA-PAINT docking sites that were imaged simultaneously with the sites of interest in all imaging rounds, thus enabling their use as fiducials. First, redundant cross-correlation (2D and 3D origami measurements) and gold alignment (3D measurements) were performed in Picasso Render. To correct for nanoscopic movement of individual DNA origami during buffer exchange, channel alignment was not only performed on the full field of view but, additionally, small regions of interest containing only one DNA origami were selected. Within each region of interest, alignment was then conducted via the fiducial docking sites of the DNA origami. This was performed outside of Picasso in a custom Python script, not only to find the optimal translation between channels but also to correct for possible rotations of the DNA origami.

Clustering and RESI

Clustering of DNA-PAINT localizations

After channel alignment, DNA-PAINT data were analysed using a custom clustering algorithm for each imaging round. This algorithm is based on the fact that, in DNA-PAINT, localizations are independent measurements of the position of a target molecule and are observed to be Gaussian distributed. To assign localizations to a specific target molecule, we first used a gradient ascent method to find the centre of a localization cloud for each target. We then assigned all localizations circularly distributed around the centre point to the same target molecule. This is a valid approximation because, due to the reduction of effective target density by RESI's sequential imaging approach, the majority of localization clouds from single targets are spaced sufficiently apart.

The clustering algorithm uses two input parameters: radius r, which sets the final size of the clusters and defines a circular environment around each localization, and the minimal number of localizations, n_{min}, representing a lower threshold for the number of DNA-PAINT localizations in any cluster.

First, the number of neighbouring localizations within distance r from each localization is calculated. If a given localization has more neighbours within its r radius than all neighbouring localizations, it is considered a local maximum. If there are more than n_{min} localizations within a circle of radius r around such a local maximum, these localizations are assigned to the same cluster; the remainder are not considered to be part of a cluster and are omitted from further analysis.

Further filtering of clusters is performed to exclude clusters that originate from unspecific sticking of imagers to the sample. Firstly, the mean frame (mean value of the frame numbers in which localizations occurred) of all localizations assigned to the same cluster is calculated. In the case of repetitive blinking the mean frame is expected to be around half the total number of frames[42]. The algorithm therefore excludes all clusters with a mean frame in the first or last 20% of frames. Secondly, sticking events in the middle of the acquisition time can be identified by dividing the acquisition time into 20 time windows each containing 5% of frames. If any of these time windows contains more than 80% of localizations in the cluster, it is excluded as a sticking event.

The choice of the clustering radius r and the threshold n_{min} depend on the respective experimental conditions. A suitable value for n_{min} can be estimated by picking localization clouds originating from single target molecules (that is, well separated) in Picasso Render, exporting pick properties and plotting a histogram of the number of localizations in each pick. n_{min} is chosen to differentiate between populations corresponding to single targets and to background localizations.

The radius r scales with the size of the localization clouds and thus the localization precision. If too large a value is chosen, adjacent clusters might not be separated; if r is too small, 'subclustering' within one localization can occur. The latter also translates to a peak in NND at twice the clustering

radius. A good a priori starting value for *r* is represented by approximately twofold the localization precision of the underlying DNA-PAINT measurement. Picasso Render offers a tool (Test Clusterer) in which the effect of different clustering parameters can be tested for a small region of interest.

For 3D clustering, an additional radius for the *z* direction is introduced because the spread of localizations in *z* is approximately twofold greater compared with *x* and *y*.

Calculation and rendering of RESI localization

Following cluster analysis, the centres of the DNA-PAINT localization groups were calculated as weighted (wtd) means by employing the squared inverse localization precisions $(\frac{1}{{\text{lp}}^{2}})$ as weights. For *x* and y coordinates:

$${\bar{x}}_{{\rm{w}}{\rm{t}}{\rm{d}}}=\frac{\mathop{\sum }\limits_{i=1}^{N}{w}_{i}{x}_{i}}{\mathop{\sum }\limits_{i=1}^{N}{w}_{i}},\,{w}_{i}=\frac{1}{{\text{lp}}^{2}}.$$

For *z* coordinates a standard mean without weights is used to calculate *z* positions. The precision of the resulting RESI localization is the weighted s.e.m. of the underlying grouped localizations:

$${({s}_{\bar{x}})}_{{\rm{wtd}}}=\frac{{({s}_{x})}_{{\rm{wtd}}}}{\sqrt{N}}=\sqrt{\frac{{\rm{Var}}{(x)}_{{\rm{wtd}}}}{N}},{\rm{where}}\,{\rm{Var}}{(x)}_{{\rm{wtd}}}=\frac{N}{N-1}\frac{{\sum }_{i=1}^{N}{w}_{i}{({x}_{i}-{\bar{x}}_{{\rm{wtd}}})}^{2}}{\mathop{\sum }\limits_{i=1}^{N}{w}_{i}}.$$

The choice for $1/{\text{lp}}^{2}$ as weights is based on the following argument: under the hypothesis that localizations are independent and normally distributed with the same mean, the weighted mean based on inverse variances as weights is the maximum likelihood estimator of the mean of the whole set of localizations. Therefore, the variance of the weighted mean is minimal (the estimator is optimal) when the inverse

variances of individual measurements $1/{{\text{lp}}}^{2}$ are chosen as weights.

Finally, we take the average of the resulting *x* and *y* s.e.m. as the final precision of each RESI localization. For *z* coordinates the precision is estimated to be two times *xy* precision. Saving RESI localizations in a Picasso hdf5 file allowed us to render them as Gaussians with s.d. corresponding to their respective precision.

RESI resolution estimation

Evaluation of in silico RESI precision with numerical simulations

To evaluate the performance of RESI, in silico numerical simulations were performed. The algorithm consists of the following steps.

1. (1)

   A grid of defined positions of the binding sites (ground truth) is generated. Typically, a grid of positions was generated (Extended Data Fig. 1a, top left).

2. (2)

   SMLM (DNA-PAINT) localizations are simulated as samples from a 2D Gaussian distribution with $\sigma = \sigma_{\text{SMLM}}$. A large number ($M$) of localizations is generated per binding site (Extended Data Fig. 1a, top right).

3. (3)

   For each binding site, subsets of K localizations are randomly selected ($K << M$). This results in $n=\frac{M}{K}$ subsets of SMLM localizations (Extended Data Fig. 1a, bottom left) that are then averaged to generate *n* RESI localizations (Extended Data Fig. 1a, bottom right).

4. (4)

The resulting RESI localizations are then shown in a histogram (Extended Data Fig. 1b) and the trace (*tr*) of the covariance matrix is calculated. RESI precision is estimated as $\sigma_{\rm{RESI}}=\sqrt{\frac{1}{2}tr({\rm{cov}}\left(x,y\right))}$ (Extended Data Fig. 1c). This definition has been used before in the field as a scalar metric for 2D variance[8].

5. (5)

   Steps 3 and 4 are repeated for different values of K to numerically study $\sigma = \sigma_{\rm{RESI}}(K)$.

Evaluation of experimental RESI precision by resampling of localizations

To evaluate the precision of RESI in experimental data, an analogous method was used. Briefly, the M total of DNA-PAINT localizations of each group corresponding to a single binding site was randomly resampled into subsets of K localizations, then steps 4 and 5 above were performed to evaluate $\sigma_{\rm{RESI}}$. The plotted $\sigma_{\rm{RESI}}$ in Fig. 3d is the average value of all single binding sites in the dataset. Error bars represent the s.d. of the different $\sigma_{\rm{RESI}}$ values calculated for different binding sites.

Note that this analysis can be performed only for $K << M$ to have sufficient $n=\frac{M}{K}$ RESI localizations for a statistically significant estimation. Because final RESI localization takes into account all M DNA-PAINT localizations, final precision is extrapolated as $\sigma_{\rm{RESI}}=\frac{\sigma_{\rm{SMLM}}}{\sqrt{M}}$.

Stochastic labelling: simulations and user guidelines

In RESI, the sparsity of binding sites in the sample is achieved by labelling a single species of biomolecules with different orthogonal DNA sequences. The labelling process is performed in a stochastic manner: n different labels (for example, DNA-conjugated nanobodies) targeting the same protein species are simultaneously incubated in the sample and thus the probability of each single protein being labelled with a certain sequence i ($i = 1, \ldots, n$) is

$p_i=\frac{1}{n}$, given that the same concentration of each label is used. Subsequently, n imaging rounds are performed to record all groups of localizations required to obtain the final RESI image.

The minimum number of labels (n) and rounds necessary to achieve sufficient sparsity of binding sites in each imaging round will depend mainly on three factors: SMLM localization precision and density and the molecular arrangement of the protein of interest. Here we describe how these parameters affect the final RESI results using a few practical examples.

Case 1: protein structure with oligomers not resolvable with DNA-PAINT

A typical study case is that of single proteins arranged in dimers, which in turn present another specific spatial organization in space. This is the case, for example, of the Nup96 in the NPC. In this case stochastic labelling has to be such that the probability of labelling two proteins forming a dimer with different sequences is sufficiently high. For n rounds of labelling/imaging, the probability is

$$P\left({\rm{diff}}\,.\,{\rm{seq}}.\right)=1-{p}_{i}=1-\frac{1}{n}$$

for $n = 4$ labelling/imaging rounds P(diff. seq.) ≈ 75%. We chose $n = 4$ to demonstrate that it provides a relatively high P(diff. seq.) with only a few imaging rounds. We note, however, that $n > 4$ could be used to increase P(diff. seq.) and hence to maximize the sparsity of labelled binding sites in each round.

To resolve a set of an arbitrary number of molecules, m, spaced more closely than the resolution of DNA-PAINT, they must be labelled with n orthogonal sequences. In general, the proportion of m molecules labelled with n orthogonal sequences, and thus the proportion of resolvable sets of molecules, follows the equation

$$P\left(m,n\right)=\frac{n!}{\left(n-m\,\right)!{n}^{m}}.$$

Case 2: proteins distributed similarly to CSR at a certain density

This is a common case—for example, for membrane receptors. If proteins are distributed in a CSR fashion (Extended Data Fig. 14a) at a given density, DNA-PAINT can already resolve single proteins that are sufficiently spaced from their NNs. We will consider that proteins at a distance $d = 4 \times \sigma_{\text{DNA-PAINT}}$ are reliably resolved (note that this criterion is significantly stricter than $2.35 \times \sigma_{\text{DNA-PAINT}}$). Then, for a given density, the NND histogram can be computed and the fraction of distances below d calculated (Extended Data Fig. 14b). This represents the fraction of single proteins, F, that will not be resolved by DNA-PAINT. Here we plot F as a function of both density and resolution (Extended Data Fig. 14c). Such a map already provides a tool to understand the level of SMLM resolution needed to resolve single proteins at a given density.

RESI can be interpreted here as a way to reduce the effective density by splitting targets into different stochastically labelled subsets. Hence, the effective density of each round will be reduced according to the formula $\rho =\frac{{\rm{density}}}{n}$. Extended Data Fig. 14d shows one-dimensional cuts of the 2D map to provide guidelines to choosing the number of orthogonal sequences (and hence imaging rounds) needed to be able to perform RESI efficiently. For example, for an initial resolution of 20 nm ($\sigma = 5$ nm), which is typical for DNA-PAINT in a cellular context, and a density of ${\rm{d}}{\rm{e}}{\rm{n}}{\rm{s}}{\rm{i}}{\rm{t}}{\rm{y}}=200\,\frac{{\rm{m}}{\rm{o}}{\rm{l}}{\rm{e}}{\rm{c}}{\rm{u}}{\rm{l}}{\rm{e}}{\rm{s}}}{\mu {{\rm{m}}}^{2}}$ (relatively high), $n = 4$ different sequences are sufficient to provide P(diff. seq.) ≈ 90% for proteins below d (Extended Data Fig. 14d). These proteins will then be resolvable by RESI.

Model-free averaging

Model-free averaging of Nup96 data was performed for both DNA-PAINT and RESI measurement of the same nucleus, as described by Wu et al.[30]. The respective Picasso hdf5 files were segmented in SMAP[43] and saved in a file format compatible for averaging by employing plugins *segmentNPC*, *NPCsegmentCleanup* and *sitenumbers2loc*. Model-free averaging was then performed on the resulting _sml.mat files with default parameters by running the particleFusion.m script in Matlab (available with the SMAP

source code). The averages shown correspond to the result of the final iteration, in which each point is rendered with a Gaussian of $\sigma = 2$ nm in x, y and z.

Numerical simulations for CD20 distribution

To interpret the results of the NND data in untreated cells, numerical simulations were performed. Briefly, two populations, one of CD20 monomers and one of dimers with a CSR distribution, were simulated and then their NNDs calculated. The algorithm can be summarized as follows:

1. (1)

   Choice of parameters. Density of monomers: number of monomers per unit area; density of dimers: number of dimers per unit area; dimer distance: expected distance between the two molecules including the labelling construct; uncertainty: variability in the position of each molecule due to labelling and localization errors; labelling efficiency: fraction of ground-truth molecules that will actually be labelled and measured. The observed density, which has to match the experimental parameter, then becomes observed density = (density of monomers + density of dimers) × labelling efficiency. For quantification of the labelling efficiency of the DNA-conjugated GFP nanobody we used a transiently transfected CHO cell line expressing a GFP- and Alfa-tag at the C terminus of a monomeric membrane protein (for example, CD86). We then labelled GFP- and Alfa-tag using their cognate nanobodies conjugated to two orthogonal docking sequences and performed two rounds of Exchange-PAINT. We then obtained the best-fitting parameters for a sample comprising pairs of GFP/Alfa-tag, and isolated Alfa-tags, similarly to how CD20 dimer/monomer analysis is performed. The ratio of these two populations is then used as an estimation of labelling efficiency. Full details of the quantification approach will be available in a manuscript currently in preparation.

2. (2)

   Simulation of monomers: a set of spatial coordinates with CSR distribution and given density are drawn; simulation of dimers: a set of

spatial coordinates with CSR distribution are drawn, representing the centre of each dimer. For each dimer centre, two positions are generated with a random orientation and a distance with expected value dimer distance. The position of each pair of molecules is drawn, taking into account the uncertainty parameter (drawn from a Gaussian distribution).

3. (3)

   A random subset of 'detectable' molecules is taken from the ground-truth set (fraction = labelling efficiency) to simulate the labelling process.

4. (4)

   NNDs are calculated on the subset of detectable molecules.

The parameters density of monomers = 212 μm^{-2}, density of dimers = 0 μm^{-2}, uncertainty = 5 nm and labelling efficiency = 50% were used to compare data for RTX-treated cells with a CSR distribution of monomers.

For the untreated case, the best-fit parameters were obtained through an iterative, nonlinear, least-squares algorithm. The experimentally observed density (50 molecules μm^{-2}) is used for the simulation.

Description of the iterative nonlinear, least-squares algorithm

For every set of parameters a simulation is performed, NNDs are histogrammed and the sum of the squared differences between the simulation and experimental histogram are computed. A fit consists of finding the parameters that minimize the sum of the squared differences.

Parameters

- D, average dimer distance (nm)
- σ_label, variability introduced by the labelling (nm)

- frac_of_dimers, fraction of dimers (%)

Note: frac_of_monomers = 100 – frac_of_dimers

Estimation of parameters

1. (1)

   Coarse-fit over a large range of parameters to determine the range of the best-fit parameters. Range D = 1–20 nm, σ_label = 1–20 nm, frac_of_dimers = 0–100%.

2. (2)

   Fine-fit over a reduced parameter space around the best-fit results in the previous step.

The parameters D_opt, σ_label_opt and frac_of_dimers_opt that best match the proposed model and the data are now found. In this case it resulted in D_opt = 13.5 nm, σ_label_opt = 5.5 nm, frac_of_dimers_opt = 47% (Fig. 4e,f).

Estimation of parameter uncertainty

1. (1)

   M is created (in this case, *M* = 100), simulated (using datasets D_opt, σ_label_opt, frac_of_dimers_opt) with the same number of molecules as the experimental data (around 21,000).

2. (2)

   M datasets are fine-fitted and the best-fit parameters D_opt, σ_label_opt and frac_of_dimers_opt are obtained. Three sets are obtained: D_opt, σ_label_opt and frac_of_dimers_opt.

3. (3)

The distributions of D_opt, σ_label_opt and frac_of_dimers_opt are studied. Standard deviation can be used as an estimation of the parameter uncertainties obtained in b.

The uncertainties of the parameters D_opt, σ_label_opt and frac_of_dimers_opt are now obtained.

Data availability

Localization data from this study are available at Zenodo (https://doi.org/10.5281/zenodo.7795826). Raw microscopy data obtained during this study are available from the corresponding author on reasonable request.

Code availability

RESI can be performed using Picasso v.0.6.0, available at https://github.com/jungmannlab/picasso with documentation provided at https://picassosr.readthedocs.io/en/latest/render.html. The custom-written scripts used in this study are available at https://github.com/jungmannlab/resi.

References

1. Hell, S. W. & Wichmann, J. Breaking the diffraction resolution limit by stimulated emission: stimulated-emission-depletion fluorescence microscopy. *Opt. Lett.* **19**, 780–782 (1994).

2. Betzig, E. Proposed method for molecular optical imaging. *Opt. Lett.* **20**, 237–239 (1995).

3. Klar, T. A., Jakobs, S., Dyba, M., Egner, A. & Hell, S. W. Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl Acad. Sci. USA* **97**, 8206–8210 (2000).

4. Betzig, E. et al. Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**, 1642–1645 (2006).

5. Rust, M. J., Bates, M. & Zhuang, X.Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nat. Methods* **3**, 793–795 (2006).

6. Sharonov, A. & Hochstrasser, R. M. Wide-field subdiffraction imaging by accumulated binding of diffusing probes. *Proc. Natl Acad. Sci. USA* **103**, 18911–18916 (2006).

7. Dai, M., Jungmann, R. & Yin, P. Optical imaging of individual biomolecules in densely packed clusters. *Nat. Nanotechnol.* **11**, 798–807 (2016).

8. Balzarotti, F. et al. Nanometer resolution imaging and tracking of fluorescent molecules with minimal photon fluxes. *Science* **355**, 606–612 (2017).

9. Schnitzbauer, J., Strauss, M. T., Schlichthaerle, T., Schueder, F. & Jungmann, R. Super-resolution microscopy with DNA-PAINT. *Nat. Protoc.* **12**, 1198–1228 (2017).

10. Reymond, L. et al. SIMPLE: structured illumination based point localization estimator with enhanced precision. *Opt. Express* **27**, 24578–24590 (2019).

11. Gu, L. et al. Molecular resolution imaging by repetitive optical selective exposure. *Nat. Methods* **16**, 1114–1118 (2019).

12. Cnossen, J. et al. Localization microscopy at doubled precision with patterned illumination. *Nat. Methods* **17**, 59–63 (2020).

13. Weber, M. et al. MINSTED fluorescence localization and nanoscopy. *Nat. Photonics* **15**, 361–366 (2021).

14. Masullo, L. A. et al. An alternative to MINFLUX that enables nanometer resolution in a confocal microscope. *Light Sci. Appl.* **11**, 199 (2022).

15. Lelek, M. et al. Single-molecule localization microscopy. *Nat. Rev. Methods Primers* https://doi.org/10.1038/s43586-021-00038-x (2021).

16. Mortensen, K. I., Churchman, L. S., Spudich, J. A. & Flyvbjerg, H. Optimized localization analysis for single-molecule tracking and super-resolution microscopy. *Nat. Methods* **7**, 377–381 (2010).

17. Jungmann, R. et al. Multiplexed 3D cellular super-resolution imaging with DNA-PAINT and Exchange-PAINT. *Nat. Methods* **11**, 313–318 (2014).

18. Jungmann, R. et al. Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Lett.* **10**, 4756–4761 (2010).

19. Rothemund, P. W. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**, 297–302 (2006).

20. Eklund, A. S., Comberlato, A., Parish, I. A., Jungmann, R. & Bastings, M. M. C. Quantification of strand accessibility in biostable DNA origami with single-staple resolution. *ACS Nano* **15**, 17668–17677 (2021).

21. Beck, M. & Hurt, E. The nuclear pore complex: understanding its function through structural insight. *Nat. Rev. Mol. Cell Biol.* **18**, 73–89 (2017).

22. Schuller, A. P. et al. The cellular environment shapes the nuclear pore complex architecture. *Nature* **598**, 667–671 (2021).

23. Loschberger, A. et al. Super-resolution imaging visualizes the eightfold symmetry of gp210 proteins around the nuclear pore complex and resolves the central channel with nanometer resolution. *J. Cell Sci.* **125**, 570–575 (2012).

24. Szymborska, A. et al. Nuclear pore scaffold structure analyzed by super-resolution microscopy and particle averaging. *Science* **341**, 655–658 (2013).

25. Heydarian, H. et al. 3D particle averaging and detection of macromolecular symmetry in localization microscopy. *Nat. Commun.* **12**, 2847 (2021).

26. Wang, W., Heydarian, H., Huijben, T., Stallinga, S. & Rieger, B. Joint registration of multiple point clouds for fast particle fusion in localization microscopy. *Bioinformatics* **38**, 3281–3287 (2022).

27. Gwosch, K. C. et al. Reply to: Assessment of 3D MINFLUX data for quantitative structural biology in cells. *Nat. Methods* **20**, 52–54 (2023).

28. Weber, M. et al. MINSTED nanoscopy enters the Angstrom localization range. *Nat. Biotechnol.* https://doi.org/10.1038/s41587-022-01519-4 (2022).

29. Agasti, S. S. et al. DNA-barcoded labeling probes for highly multiplexed Exchange-PAINT imaging. *Chem. Sci.* **8**, 3080–3091 (2017).

30. Wu, Y. L. et al. Maximum-likelihood model fitting for quantitative analysis of SMLM data. *Nat. Methods* **20**, 139–148 (2023).

31. Drew, H. R. et al. Structure of a B-DNA dodecamer: conformation and dynamics. *Proc. Natl Acad. Sci. USA* **78**, 2179–2183 (1981).

32. Strauss, S. & Jungmann, R. Up to 100-fold speed-up and multiplexing in optimized DNA-PAINT. *Nat. Methods* **17**, 789–791 (2020).

33. Helmerich, D. A. et al. Photoswitching fingerprint analysis bypasses the 10-nm resolution barrier. *Nat. Methods* **19**, 986–994 (2022).

34. Pavlasova, G. & Mraz, M. The regulation and function of CD20: an "enigma" of B-cell biology and targeted therapy. *Haematologica* **105**, 1494–1506 (2020).

35. Glennie, M. J., French, R. R., Cragg, M. S. & Taylor, R. P. Mechanisms of killing by anti-CD20 monoclonal antibodies. *Mol. Immunol.* **44**, 3823–3837 (2007).

36. Pierpont, T. M., Limper, C. B. & Richards, K. L. Past, present, and future of rituximab—the world's first oncology monoclonal antibody therapy. *Front. Oncol.* **8**, 163 (2018).

37. Rouge, L. et al. Structure of CD20 in complex with the therapeutic monoclonal antibody rituximab. *Science* **367**, 1224–1230 (2020).

38. Kumar, A., Planchais, C., Fronzes, R., Mouquet, H. & Reyes, N. Binding mechanisms of therapeutic antibodies to human CD20. *Science* **369**, 793–799 (2020).

39. Beliu, G. et al. Bioorthogonal labeling with tetrazine-dyes for super-resolution microscopy. *Commun. Biol.* **2**, 261 (2019).

40. Douglas, S. M. et al. Rapid prototyping of 3D DNA-origami shapes with caDNAno. *Nucleic Acids Res.* **37**, 5001–5006 (2009).

41. Edelstein, A. D. et al. Advanced methods of microscope control using muManager software. *J. Biol. Methods* **1**, e10 (2014).

42. Wade, O. K. et al. 124-Color super-resolution imaging by engineering DNA-PAINT blinking kinetics. *Nano Lett.* **19**, 2641–2646 (2019).

43. Ries, J. SMAP: a modular super-resolution microscopy analysis platform for SMLM data. *Nat. Methods* **17**, 870–872 (2020).

Acknowledgements

We thank Y.-L. Wu and J. Ries for valuable assistance with model-free averaging. We thank J. Schmied and F. Schueder for helpful discussions. We thank M. K. Steen-Mueller and I. Glueck for proofreading the manuscript. This research was funded in part by the European Research Council through an ERC Consolidator Grant (ReceptorPAINT, grant agreement no. 101003275), the German Research Foundation through SFB1032 (project A11, no. 201269156), the Danish National Research Foundation (Centre for Cellular Signal Patterns, DNRF135), the Human Frontier Science Program through a Young Investigator Grant (no. HFSP RGY0065/2018), the Volkswagen Foundation through the initiative 'Life?—A Fresh Scientific

Approach to the Basic Principles of Life’ (grant no. 98198), the Max Planck Foundation and the Max Planck Society. S.S. and T.S. acknowledge support by the QBM graduate school. S.C.M.R., I.B., P.R.S., A.S.E., E.M.U. and M.T.S. acknowledge support by the IMPRS-LS graduate school. I.B. acknowledges funding support by Roche. L.A.M. acknowledges a postdoctoral fellowship from the European Union’s Horizon 20212022 research and innovation programme under Marie Skłodowska-Curie grant agreement no. 101065980.

Funding

Open access funding provided by the Max Planck Society.

Author information

Author notes

1. These authors contributed equally: Susanne C. M. Reinhardt, Luciano A. Masullo, Isabelle Baudrexel, Philipp R. Steen

Authors and Affiliations

1. Max Planck Institute of Biochemistry, Planegg, Germany

   Susanne C. M. Reinhardt, Luciano A. Masullo, Isabelle Baudrexel, Philipp R. Steen, Rafal Kowalewski, Alexandra S. Eklund, Sebastian Strauss, Eduard M. Unterauer, Thomas Schlichthaerle, Maximilian T. Strauss & Ralf Jungmann

2. Faculty of Physics and Center for NanoScience, Ludwig Maximilian University, Munich, Germany

   Susanne C. M. Reinhardt, Philipp R. Steen, Rafal Kowalewski, Sebastian Strauss, Eduard M. Unterauer, Thomas Schlichthaerle, Maximilian T. Strauss & Ralf Jungmann

3. Department of Chemistry and Biochemistry, Ludwig Maximilian University, Munich, Germany

Isabelle Baudrexel, Alexandra S. Eklund & Christian Klein

4. Roche Innovation Center Zurich, Roche Pharma Research and Early Development, Schlieren, Switzerland

Christian Klein

Contributions

S.C.M.R. designed and conducted 2D and 3D DNA origami as well as Nup96 experiments, developed the analysis software and analysed DNA origami and Nup96 data. L.A.M. designed and conducted computer simulations, contributed to analysis software and analysed DNA origami, Nup96 and CD20 data. I.B. designed and conducted CD20 experiments and analysed CD20 data. P.R.S. designed and conducted 2D DNA origami experiments, contributed to analysis software and analysed DNA origami and Nup96 data. R.K. and T.S. contributed to analysis software. S.S. developed labelling probes. A.S.E., S.S., E.M.U. and M.T.S. performed preliminary RESI experiments. C.K. contributed to the design of studies targeting CD20 and their interpretation. S.C.M.R., L.A.M., I.B., P.R.S. and R.J. interpreted data and wrote the manuscript. R.J. conceived the concept, designed experiments and supervised the study. S.C.M.R., L.A.M., I.B. and P.R.S. contributed equally. All authors reviewed and approved the final manuscript.

Corresponding author

Correspondence to Ralf Jungmann.

Ethics declarations

Competing interests

C.K. declares employment, patents (unrelated to this work) and stock ownership with Roche.

Peer review

Peer review information

Nature thanks Alistair Curd and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 RESI resolution estimation.

a, A grid of defined positions of binding sites is generated (top left), SMLM (DNA-PAINT) localizations are simulated as samples from a gaussian distribution (top right). Localizations for only one binding site were plotted for clarity. For each binding site, subsets of K localizations are randomly selected (bottom left) and averaged (bottom right). One exemplary subset and its average is highlighted. **b**, Resulting RESI-localizations are histogrammed to produce images at different resolutions (K values). **c**, RESI-localization precision σ_{RESI} vs K. Analytical dependence on \(\sqrt{K}\) (blue line) and numerical results (black dots). A total of 1200 SMLM localizations per site are simulated. Error bars represent mean ± 1 s.d.

Extended Data Fig. 2 RESI in 2D DNA origami.

a, DNA origami design featuring six 5 nm-spaced orthogonal docking strand pairs (red R1, blue R3) and six alignment docking strands (green R4). See [Methods](#) for sequence details. **b**, DNA-PAINT acquisition parameters were tuned such that 5 nm were not consistently resolvable. **c**, First imaging round conducted with R1 (target) and R4 imagers (alignment, sites circled). **d**, Second imaging round conducted with R3 (target) and alignment imagers (R4, sites circled). The alignment sites were used for translational and rotational alignment between rounds. **e**, RESI resolves the 5 nm distances. **f**, The distance and orientation between R1 and R3 docking strands are consistent with the design. **g**, An average of 90 DNA origami structures reveals consistent results and excellent alignment performance. The numbers indicate the distance between rounds.

Extended Data Fig. 3 2D DNA origami.

Representative DNA origami from across the field of view of the measurement. **a**, Four DNA origami, shown at DNA-PAINT resolution (upper row) and RESI resolution (lower row). The insert depicts a pair of docking strands spaced at approx. 5 nm. **b**, 40 additional DNA origami, shown at DNA-PAINT resolution (upper rows) and RESI resolution (lower rows).

Extended Data Fig. 4 RESI in 3D DNA origami.

a, DNA origami design featuring one pair of orthogonal docking strands (red R1, blue R3) as well as six alignment docking strands (green R4). Docking strands extend from both the top and bottom surface of the DNA origami (insert). **b**, The design ensures that all but the R1/R3 docking strand pair are spaced sufficiently to be resolved by DNA-PAINT. **c**, 3D DNA-PAINT imaging resolves R4 alignment sites, barely resolves R1/R3 axially and does not resolve R1/R3 laterally. **d**, Sequential 3D DNA-PAINT imaging with R4 sites used for alignment. **e**, RESI resolves R1/R3 both axially and laterally. **f**, An overlay of 88 DNA origami reveals overall good alignment despite structural heterogeneity. **g**, Average of 88 DNA origamis. **h**, The particle average recovers the structure with an alignment uncertainty of 0.7 nm CI = [0, 1.6] nm, showing a distance between the average R1/R3 positions of 11.6 ± 0.8 nm (xy-distance: 2.5 ± 0.4 nm, z-distance: 11.3 ± 0.8 nm),

matching the designed distances[20]. Same scale applies to all magnification panels. CI describes 68% confidence interval.

Extended Data Fig. 5 3D DNA origami.

Representative 3D DNA origami from across the field of view of the measurement. **a**, Five DNA origami, shown at DNA-PAINT resolution (upper row) and RESI resolution (lower row). The color scale to the right represents the z position of localizations. The measured z coordinates for each DNA origami have been shifted by a constant such that the lowest localization for a given structure is defined to be at z = 0. This ensures full use of the color range. **b**, 32 additional DNA origami, shown at DNA-PAINT resolution (upper rows) and RESI resolution (lower rows). The z positions are colored according to the color scale in panel **a**.

Extended Data Fig. 6 U2OS Nup96-mEGFP.

Representative NPCs from across the field of view of the measurement. **a**, Six NPCs, measured using DNA-PAINT (upper row) and RESI (lower row). The color scale to the right represents the z position of localizations. The measured z coordinates for each NPC have been shifted by a constant such that the lowest localization for a given structure is defined to be at z = 0. This ensures full use of the color range. **b**, 72 additional NPCs, measured using DNA-PAINT (upper rows) and RESI (lower rows). The z positions are colored according to the color scale in panel **a**.

Extended Data Fig. 7 Averaging of Nup96 proteins.

a, Model-free averaging for DNA-PAINT measurements of Nup96 (N = 1045 NPCs). An angled isometric view is shown. **b**–**d**, DNA-PAINT resolves nucleoplasmic and cytoplasmic rings and recapitulates their eight-fold symmetry, but fails to resolve individual Nup96 proteins. **e**, Side views of all Nup96 pairs in both rings reveal the angled orientation but do not resolve individual Nup96 proteins. **f**, Model-free averaging for RESI measurements of Nup96 (N = 1190 NPCs). **g**–**i**, RESI recapitulates nucleoplasmic and cytoplasmic rings as well as their eight-fold symmetry and resolves individual adjacent Nup96 proteins in the majority of cases. **j**,

Side views of all eight Nup96 pairs in both rings reveal the angled orientation as well as, in some cases, adjacent individual Nup96 proteins. **k**, The Cryo-EM structure of the nuclear pore complex indicates that a given Nup96 protein will have neighbors spaced at 11 nm, 39 nm, 71 nm, 93 nm and 101 nm. **l**, Performing clustering and nearest neighbor analysis for DNA-PAINT data reveals a peak at approx. 40 nm, corresponding to the distance between two Nup96 pairs, but not below that. RESI, on the other hand, features a first peak at approx. 15 nm, corresponding to the distance between adjacent Nup96 while taking linkage error (label size) into account. **m**, Analysis of first to tenth nearest neighbor distances for RESI and DNA-PAINT recapitulates the distances from (**k**), but only RESI resolves the smallest distance. All scale bars: 20 nm.

Extended Data Fig. 8 Sub-nm DNA origami.

Representative DNA origami from across the field of view of the measurement. **a**, Four DNA origami, shown at DNA-PAINT resolution (upper row) and RESI resolution (lower row). The inserts show pairs of directly adjacent docking strands resolved by RESI. **b**, 42 additional DNA origami, shown at DNA-PAINT resolution (upper rows) and RESI resolution (lower rows).

Extended Data Fig. 9 Sub-nm RESI measurements.

a, DNA origami featuring six alignment strands (green R4) and six pairs of orthogonal docking strands (red R1, blue R3) spaced one base pair apart. **b**, RESI representation with RESI-localizations from round 1 in red and round 2 in blue illustrates excellent alignment. The distances between RESI-localizations from round 1 and 2 are defined as illustrated. **c**, Overlaying 42 DNA origami and performing a particle average recovers the structure with an alignment uncertainty of 1.2 Å CI = [0, 4.6] Å, showing distances between the average positions of the sites at 9.5 ± 2.6 Å (mean over six distances in the average ± mean over the error-propagated uncertainties of the six distances). Same scale applies to all magnification panels. CI describes 68% confidence interval.

Extended Data Fig. 10 RESI resolves CD20 dimers in untreated CHO cells for different expression levels.

a, DNA-PAINT imaging of whole mEGFP-CD20-expressing CHO cells, labeled with anti-GFP-nanobodies, shows homogeneously distributed molecules for three independent experiments. **b**, Zoom-in regions of DNA-PAINT show cases in which dimers could not be resolved. **c**, RESI reveals sub-10-nm spaced receptor pairs, which are unresolvable in the DNA-PAINT cases. **d**, Whole-cell analysis of first nearest neighbor distances (1st NNDs) of CD20 receptors (histograms of the distances are displayed). Only RESI, but not DNA-PAINT, allows the routine detection of sub-10-nm distances between proteins. **e**, RESI-localization precision below 1 nm allows for routine detection of sub-10-nm distances, resulting in an accurate assessment of the first NND.

Extended Data Fig. 11 RESI resolves the substructure in RTX-induced chain-like arrangements of CD20 receptors with sub-nanometer precision.

a, DNA-PAINT overview image of mEGFP-CD20 expressing CHO cells treated with RTX. **b**, Labeling with DNA-conjugated anti-GFP-nanobodies and imaging with DNA-PAINT reveals higher-order organization after RTX-treatment. RESI (insets i–iii) achieves molecular resolution and thereby resolves the molecular arrangement of mEGFP-CD20. **c**, DNA-PAINT imaging shows clustered CD20 molecules. Performing RESI with sequences R1, R2, R3 and R4 in four separate imaging rounds (color-coded) allows for clustering of localizations originating from a single target. From the clustered localizations, RESI-localizations were calculated, enabling true single-protein resolution.

Extended Data Fig. 12 RESI reveals higher order arrangement of CD20 dimers in Rituximab-treated CHO cells.

a, DNA-PAINT imaging of whole mEGFP-CD20-expressing CHO cells, labeled with anti-GFP-nanobodies, shows clustered CD20- molecules in Rituximab-treated cells for three independent experiments. **b**, Zoom-in

regions of DNA-PAINT show mEGFP- CD20 clustered into chain-like arrangements. **c**, RESI reveals sub-10-nm spaced receptor pairs within the clusters, unresolvable by DNA-PAINT. **d**, Whole-cell analysis of first nearest neighbor distances (1^{st} NNDs) of CD20 receptors bound to Rituximab (histograms of the distances are displayed). Only RESI, but not DNA-PAINT, allows the routine detection of sub-10-nm distances between proteins. **e**, Routine detection of sub-10-nm distances by RESI recapitulates the first NND measured in the untreated case. Notably the NND peaks measured in the three repeats are consistent, independently of the protein density.

Extended Data Fig. 13 Comparison of Rituximab treated CD20 data to simulated CD20 hexamer arrangements.

a, Example of ground truth simulated CD20 hexamers (light blue circles, simulated as triangles of dimers with intra-dimer distances of 13.5 nm as measured experimentally) with random distribution and orientation on a 2D surface at the experimentally determined density. **b**, Label uncertainty and labeling efficiency (black circles indicate labeled molecules) are taken into account in the simulation for a realistic comparison. **c**, Simulated proteins in hexameric arrangements represented as gaussians. **d**, Hexamers after DBSCAN cluster analysis (colors indicate clusters). **e**, RESI image of CD20 data after RTX-treatment. **f**, RESI-localizations of CD20 data after DBSCAN cluster analysis (colors indicate clusters). **g**, Number of molecules per detected cluster for the experimental data and the simulated hexamers. **h**, Circularity metric of experimental data and the simulated hexamers after convex hull analysis of the clusters. We note that the sharp drop at 0.605 stems from the maximum circularity metric for clusters where the convex hull is defined by three molecules. Notably, the absence of a circularity peak at ~0.7 in the experimental data suggests that CD20 molecules are not arranged in isolated ring-like hexameric structures.

Extended Data Fig. 14 Stochastic labeling.

a, Exemplary simulation of proteins with a Complete Spatial Random (CSR) distribution of a given density. **b**, Histogram of Nearest Neighbor Distances (NNDs). The red line indicates the smallest distance (*d*) that can be resolved

by DNA-PAINT for a given set of imaging parameters. The fraction of molecules with a NN below this distance threshold (blue, shaded) can be computed for a given density and a given DNA-PAINT resolution. **c**, 2D map of the fraction of non-resolvable molecules as a function of density and resolution. **d**, 1D cuts of c at different resolutions (color-coded) can be used as a user guide to estimate the number of multiplexing rounds needed to perform RESI efficiently given a certain target fraction of non-resolvable distances.

Extended Data Table 1 Imaging and RESI parameters

Supplementary information

Supplementary Data 1

Staple-strand sequences for 2D DNA origami with subnanometre-, 5-nm- and 20-nm-spaced binding sites and staple-strand sequences for 3D origami.

Supplementary Data 2

Scaffold-strand sequence (M13mp18) for 2D DNA origami.

Supplementary Data 3

Scaffold-strand sequence (p7560) for 3D DNA origami.

Peer Review File

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's

Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Reinhardt, S.C.M., Masullo, L.A., Baudrexel, I. *et al.* Ångström-resolution fluorescence microscopy. *Nature* **617**, 711–716 (2023). https://doi.org/10.1038/s41586-023-05925-9

- Received: 21 July 2022
- Accepted: 07 March 2023
- Published: 24 May 2023
- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-05925-9

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

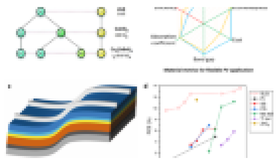
Provided by the Springer Nature SharedIt content-sharing initiative

Cocktails of tags enhance resolution of microscopy technique

- Alistair Curd

Nature News & Views 24 May 2023

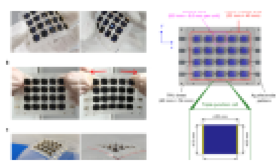
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05925-9



Emergence of flexible kesterite solar cells: progress and perspectives

23 March 2023

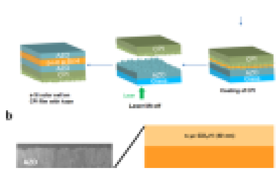
Jianjun Li, Kaiwen Sun, … Xiaojing Hao



Stretchable micro-scale concentrator photovoltaic module with 15.4% efficiency for three-dimensional curved surfaces

08 January 2021

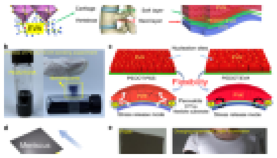
Daisuke Sato, Taizo Masuda, … Noboru Yamada



Flexible and transparent thin-film light-scattering photovoltaics about fabrication and optimization for bifacial operation

27 March 2023

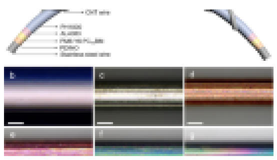
Soo-Won Choi, Jae-Ho Park, … Jung-Dae Kwon



[Bio-inspired vertebral design for scalable and flexible perovskite solar cells]

15 June 2020

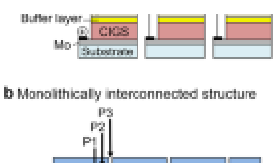
Xiangchuan Meng, Zheren Cai, … Yiwang Chen



[Highly efficient fiber-shaped organic solar cells toward wearable flexible electronics]

07 June 2022

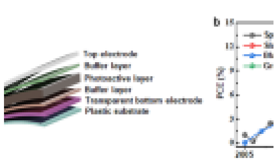
Dan Lv, Qianqing Jiang, … Dianyi Liu



[Lightweight and flexible Cu(In,Ga)Se2 solar minimodules: toward 20% photovoltaic efficiency and beyond]

30 October 2022

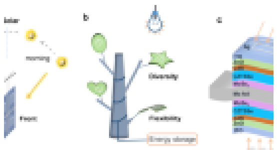
Shogo Ishizuka, Yukiko Kamikawa & Jiro Nishinaga



[Large-area flexible organic solar cells](#)

04 November 2021

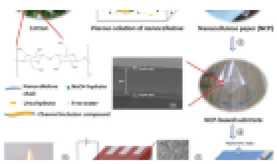
Fu Yang, Yuting Huang, … Yongfang Li



[Novel symmetrical bifacial flexible CZTSSe thin film solar cells for indoor photovoltaic applications](#)

25 May 2021

Hui Deng, Quanzhen Sun, … Shuying Cheng



[Flexible, transparent nanocellulose paper-based perovskite solar cells](#)

11 February 2019

Lei Gao, Lingfeng Chao, … Wei Huang

- Article
- Open Access
- Published: 24 May 2023

Flexible solar cells based on foldable silicon wafers with blunted

edges

- Wenzhu Liu ,
- Yujing Liu,
- Ziqiang Yang,
- Changqing Xu,
- Xiaodong Li,
- Shenglei Huang,
- Jianhua Shi,
- Junling Du,
- Anjun Han,
- Yuhao Yang,
- Guoning Xu,
- Jian Yu,
- Jiajia Ling,
- Jun Peng,
- Liping Yu,
- Bin Ding,
- Yuan Gao,
- Kai Jiang,
- Zhenfei Li,
- Yanchu Yang,
- Zhaojie Li,
- Shihu Lan,
- Haoxin Fu,
- Bin Fan,
- Yanyan Fu,
- Wei He,
- Fengrong Li,
- Xin Song,
- Yinuo Zhou,
- Qiang Shi,
- Guangyuan Wang,
- Lan Guo,
- Jingxuan Kang,
- Xinbo Yang,

- Dongdong Li,
- Zhechao Wang,
- Jie Li,
- Sigurdur Thoroddsen,
- Rong Cai,
- Fuhai Wei,
- Guoqiang Xing,
- Yi Xie,
- Xiaochun Liu ,
- Liping Zhang ,
- Fanying Meng ,
- Zengfeng Di &
- …
- Zhengxin Liu

Show authors

Nature **volume 617**, pages 717–723 (2023)

- 10k Accesses
- 74 Altmetric
- Metrics details

Abstract

Flexible solar cells have a lot of market potential for application in photovoltaics integrated into buildings and wearable electronics because they are lightweight, shockproof and self-powered. Silicon solar cells have been successfully used in large power plants. However, despite the efforts made for more than 50 years, there has been no notable progress in the development of flexible silicon solar cells because of their rigidity[1,2,3,4]. Here we provide a strategy for fabricating large-scale, foldable silicon wafers and manufacturing flexible solar cells. A textured crystalline silicon wafer always starts to crack at the sharp channels between surface pyramids in the marginal region of the wafer. This fact enabled us to improve the

flexibility of silicon wafers by blunting the pyramidal structure in the marginal regions. This edge-blunting technique enables commercial production of large-scale (>240 cm^2), high-efficiency (>24%) silicon solar cells that can be rolled similarly to a sheet of paper. The cells retain 100% of their power conversion efficiency after 1,000 side-to-side bending cycles. After being assembled into large (>10,000 cm^2) flexible modules, these cells retain 99.62% of their power after thermal cycling between −70 °C and 85 °C for 120 h. Furthermore, they retain 96.03% of their power after 20 min of exposure to air flow when attached to a soft gasbag, which models wind blowing during a violent storm.

Main

Silicon is the most abundant semiconducting element in Earth's crust; it is made into wafers to manufacture approximately 95% of the solar cells in the current photovoltaic market[5]. However, these cells are brittle and crack under bending stress, which limits their large-scale use for flexible applications. At present, thin-film solar cells made from amorphous silicon, $Cu(In,Ga)Se_2$, CdTe, organics and perovskites exhibit flexibility[6,7,8,9] but their use is limited because of their low power conversion efficiency (PCE), release of toxic materials into the environment, inferior performance in the case of large areas and unstable operating conditions. Therefore, many available flexible solar cells have not attracted customers and most companies that manufactured them have gone out of business. In this study, we propose a morphology engineering method to fabricate foldable crystalline silicon (c-Si) wafers for large-scale commercial production of solar cells with remarkable efficiency.

Fabrication of foldable c-Si wafers

Our first goal was to fabricate foldable c-Si wafers with a strong light-harvesting ability. Reducing the thickness of a wafer can improve its flexibility[10], but there is a trade-off between thickness and light-harvesting efficiency because c-Si is a semiconductor with an indirect optical bandgap. By using saw-damage removal[11], we reduced the thickness of a 160-μm wafer to 60 μm. Although the wafer began to exhibit flexibility similar to

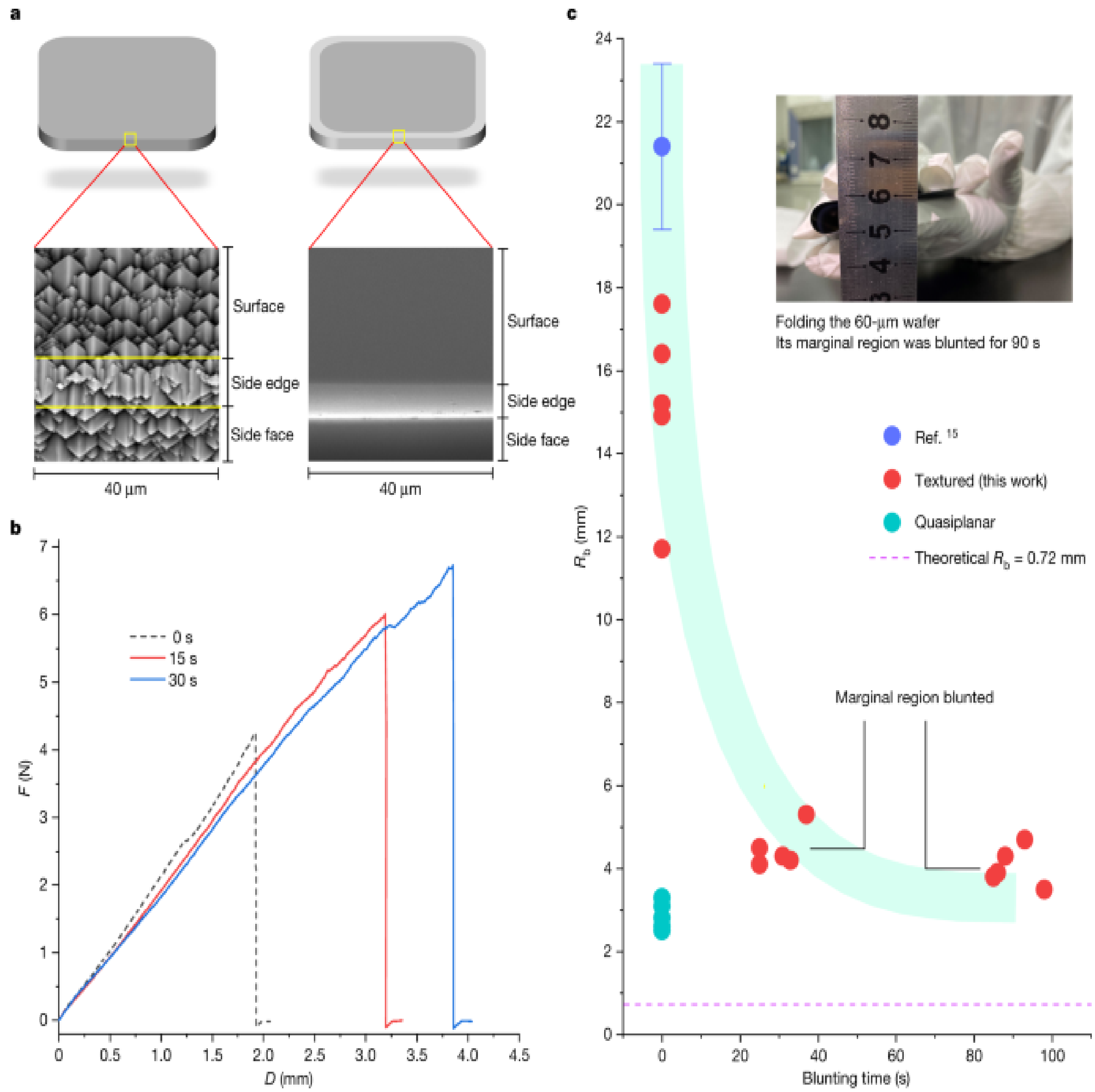
that of a sheet of paper (Supplementary Fig. 1), it was not suitable for solar cell fabrication because more than 30% of the incident sunlight was reflected by its glossy surface[12]. Chemically texturing microscale pyramids on c-Si surfaces has been widely used as an efficient strategy to reduce the reflectivity to less than 10% owing to Lambertian light trapping[13]. However, when bending forces were applied to such textured wafers, the maximum stress was located in the sharp channels between the pyramids, as observed in the simulation with the solid mechanics module in COMSOL Multiphysics (Extended Data Fig. 1a). This result was consistent with an in situ image obtained using transmission electron microscopy (TEM), in which the bending stress accumulated in the channels between pyramids under a typical bending load exerted by a micromanipulator (Extended Data Fig. 2). Further simulations revealed that a slight increase in the channel radius (R_p) from 0 μm to 2.3 μm led to a rapid reduction in the maximum stress from 0.25 MPa to 0.016 MPa (Extended Data Fig. 1b). But this blunting treatment increased the reflectivity to more than 30% (Supplementary Fig. 2), which was unfavourable for light harvesting. This was confirmed by optical simulations of the devices, in which the blunted wafers showed inferior antireflection and light trapping (Supplementary Fig. 3).

Next, we used an ultrahigh-speed video camera to investigate the cracking process of a wafer. As shown in Extended Data Fig. 3, the camera recorded a long fracture in the snapshots taken at 113, 132 and 151 μs (yellow arrows). If we assume that the cracking initiates from a point on the edge (circle), then the propagation velocities of the crack can be estimated from the evolution of the fracture length as 33.2, 33.6 and 33.0 m s^{-1}, respectively. The consistency of these values supported our hypothesis. Furthermore, the camera recorded three silicon particles being ejected from the edge of the wafer (red arrows); their initial positions coincided with the point at which the cracking initiated. This extra evidence confirmed that the cracking initiated at the edge of the wafer, which explains why most linear cracks in electroluminescence images begin from the edges of silicon solar cells[14] (Supplementary Fig. 4).

On the basis of the cracking characteristics discussed above, we considered blunting the sharp channels in the marginal region instead of the whole

wafer to improve the flexibility of the silicon wafers (Fig. 1a). The three-point bending test results in Fig. 1b show that the 15-s and 30-s edge-blunting treatments increased the vertical displacement of the wafer at the cracking moment from 1.92 mm to 3.20 mm and 3.86 mm, respectively. Consistent with these results, after the approximately 2-mm-wide marginal region of a 60-μm textured wafer was blunted for 0, 30 and 90 s in 10 vol% HF:90 vol% HNO_3 solution, the critical bending radius (R_b) at the cracking moment considerably decreased from 15.2 ± 2 mm (this study) or 21.4 ± 2 mm (ref. [15]) to approximately 4.0 mm, which approached the theoretical limit of 0.72 mm (Fig. 1c). As expected, we could fold the wafer around its centre with $R_b = 4.0$ mm (Fig. 1c, inset) and vigorously shake it like a sheet of flexible paper (Supplementary Video 1). The improvement in flexibility was also supported by atomistic simulations: cracking of the untreated wafer initiated under a loading strain of 9.3%, but this value rapidly increased to 17.3% (Supplementary Videos 2–5) even when the sharp channels were blunted to a very small R_p of 15.81 nm.

Fig. 1: Foldable wafers.



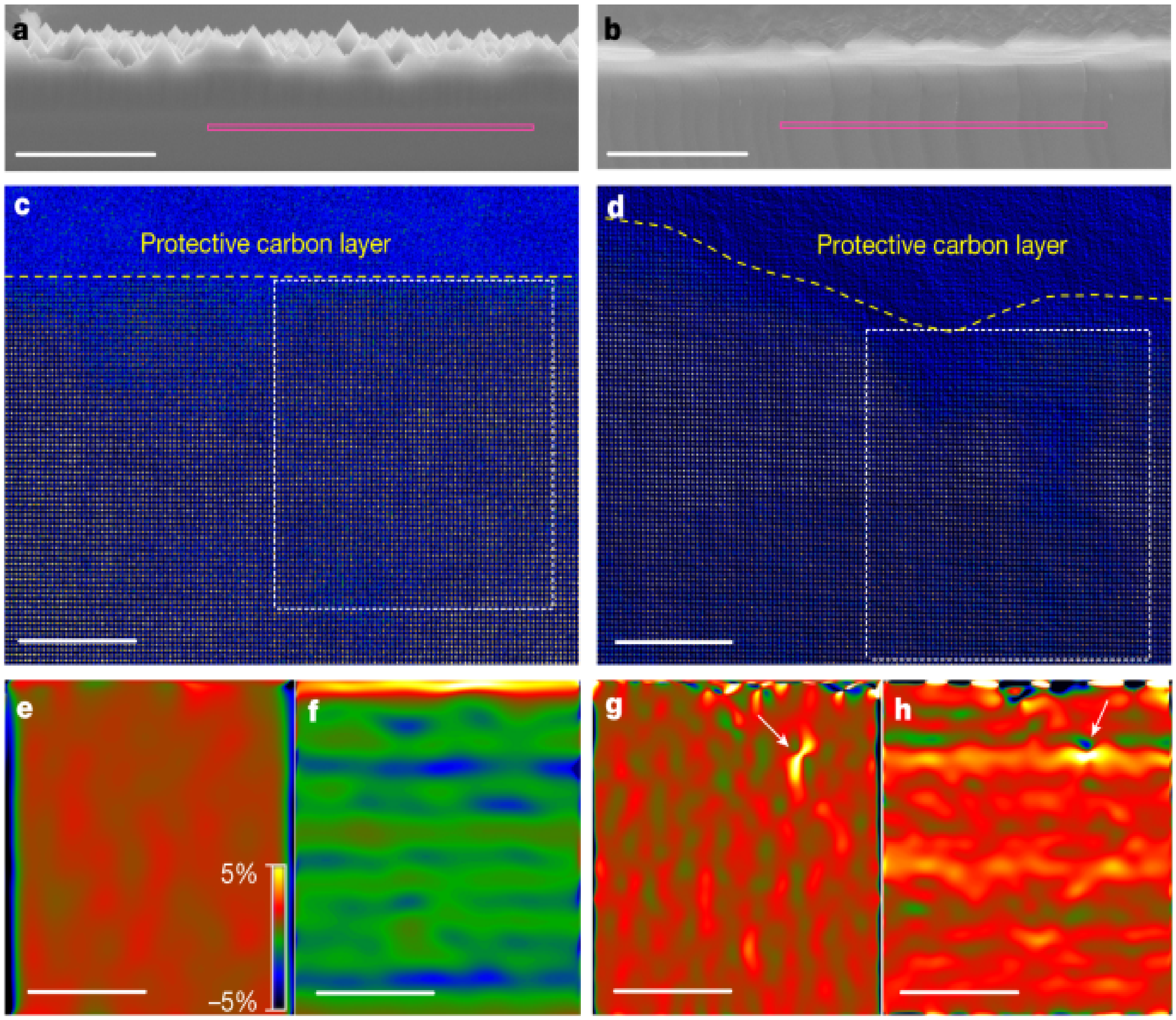
a, SEM images of a textured c-Si wafer. The sharp pyramids in the marginal region were efficiently removed by an acid solution. **b**, Load–vertical displacement (*F*–*D*) curves of 140-µm textured c-Si wafers, in which the marginal regions were blunted in 10 vol% HF:90 vol% HNO_3 solution for 0, 15 and 30 s. **c**, Textured c-Si wafers (60 µm) with pyramids on the surface. Their bending radius (R_b) at the cracking moment is plotted as a function of the blunting time in 10 vol% HF:90 vol% HNO_3 solution. For comparison, the R_b values of 60-µm textured[15] and quasiplanar c-Si wafers are shown. We also calculated the theoretical R_b of a 60-µm c-Si wafer as $R_b = Ed/2\sigma$, where E, d and σ are the modulus of elasticity, wafer thickness and tensile

yield strength, respectively. Inset, 60-μm textured wafer of size 15.6 cm × 15.6 cm, in which the marginal region was blunted for 90 s in the acid solution.

Flexibility mechanism

To understand the flexibility of the c-Si wafers shown in Fig. 1c, we broke two wafers by applying bending forces to find out the morphologies of the fracture surfaces. The image obtained by scanning electron microscopy (SEM) of the untreated wafer in Fig. 2a shows a flat cleavage surface, whereas the SEM image of the blunted wafer in Fig. 2b shows a fracture surface with multiple cleavage sites and a high density of microcracks, which is also evident in the stepwise focused ion beam (FIB) image of the fracture surface (Supplementary Fig. 5). In the magnified view of the fracture surface with multiple cleavage sites (Extended Data Fig. 4), we observed large cracks propagating along complicated paths in deeper regions of the blunted wafer (yellow arrows) and some jagged notches (pink arrows), which were in good agreement with the atomistic simulations (Extended Data Fig. 5). Within a depth of around 500 nm below the top surface (white arrows), secondary shear banding lines (red arrows) were generated in a direction different from that of the dominant cracks (yellow arrows; Extended Data Fig. 4). These features indicated the development of a complex stress state during the cracking process, which was similar to the restoration of deformability of brittle metallic glass realized by triggering secondary shear banding[16]. Physically, these bumpy cleavage processes consumed more energy before the initiation of cracking; thus, they accounted for the robust behaviour that provided protection against the violent folding (Fig. 1c).

Fig. 2: GPA of the fracture surfaces.



a,**b**, Morphology of the fracture surface of a wafer with sharp (**a**) and round (**b**) pyramids. The pink lines mark the locations at which the top surface of the fracture was protected and lifted out for TEM observations using an FIB. **c,d**, High-resolution STEM-HAADF images showing the atomic arrangement at a depth of dozens of atoms viewed along the [001] direction from the fracture surface of the wafer with sharp (**c**) and round (**d**) pyramids, in which a protective carbon layer was deposited on the fracture surface. The GPA regions are highlighted by dashed squares. **e**,**f**, Elastic lattice strain distribution in the *x* direction (**e**) and *y* direction (**f**) of the wafer with sharp pyramids. **g**,**h**, Elastic lattice strain distribution in the *x* direction (**g**) and *y* direction (**h**) of the wafer with round pyramids. White arrows mark the great dilatation strain. Positive and negative values represent lattice expansion and contraction, respectively. The *x* direction is parallel and the *y* direction is

perpendicular to the fracture surfaces marked in **a** and **b**, respectively. Scale bars, 5 μm (**a**,**b**); 5 nm (**c**–**h**).

Using spherical aberration-corrected TEM, we analysed the lattice strains beneath the fracture surface. After depositing a protective carbon layer on the fresh fracture surface, we obtained atomic-resolution images using high-angle annular dark field scanning transmission electron microscopy (HAADF-STEM) of the untreated (Fig. 2c) and blunted (Fig. 2d) wafers. The rough fracture surface marked by a yellow dashed line in Fig. 2d indicates that the blunted wafer underwent more elastic and plastic strain during the cracking process. Because some of the lattice strain caused by lattice distortion could be preserved in dozens of atomic layers beneath the fracture surface, we could analyse the residual strain as an indicator of the cracking mode. Figure 2e,f shows the geometric phase analyses[17] (GPAs) of the area beneath the fracture surface of the untreated wafer (Fig. 2c). This wafer showed general tensile strain in the *x* direction and compressive strain in the *y* direction but showed remarkable dilatation strain in the *y* direction within a few atomic layers of the top surface. These characteristics pertained to a typical brittle fracture mode. By contrast, Fig. 2g,h shows that the fracture surface of the blunted wafer had larger lattice strain variations in both the *x* and *y* directions; the great dilatation strain is marked by white arrows. This feature suggests that complex cracking can result in much larger lattice expansion. These findings prove that the fracture behaviour of c-Si wafers can be manipulated by tuning the sharpness of the channels between surface pyramids, which modifies the stress state and deformation mechanism under bending loads. As a consequence, in this study, the blunting treatment greatly mitigated the intrinsic brittleness of the c-Si wafer, which led to a transition of the fracture mechanism from intrinsic brittle cleavage fracture to shear banding with steps and cracks.

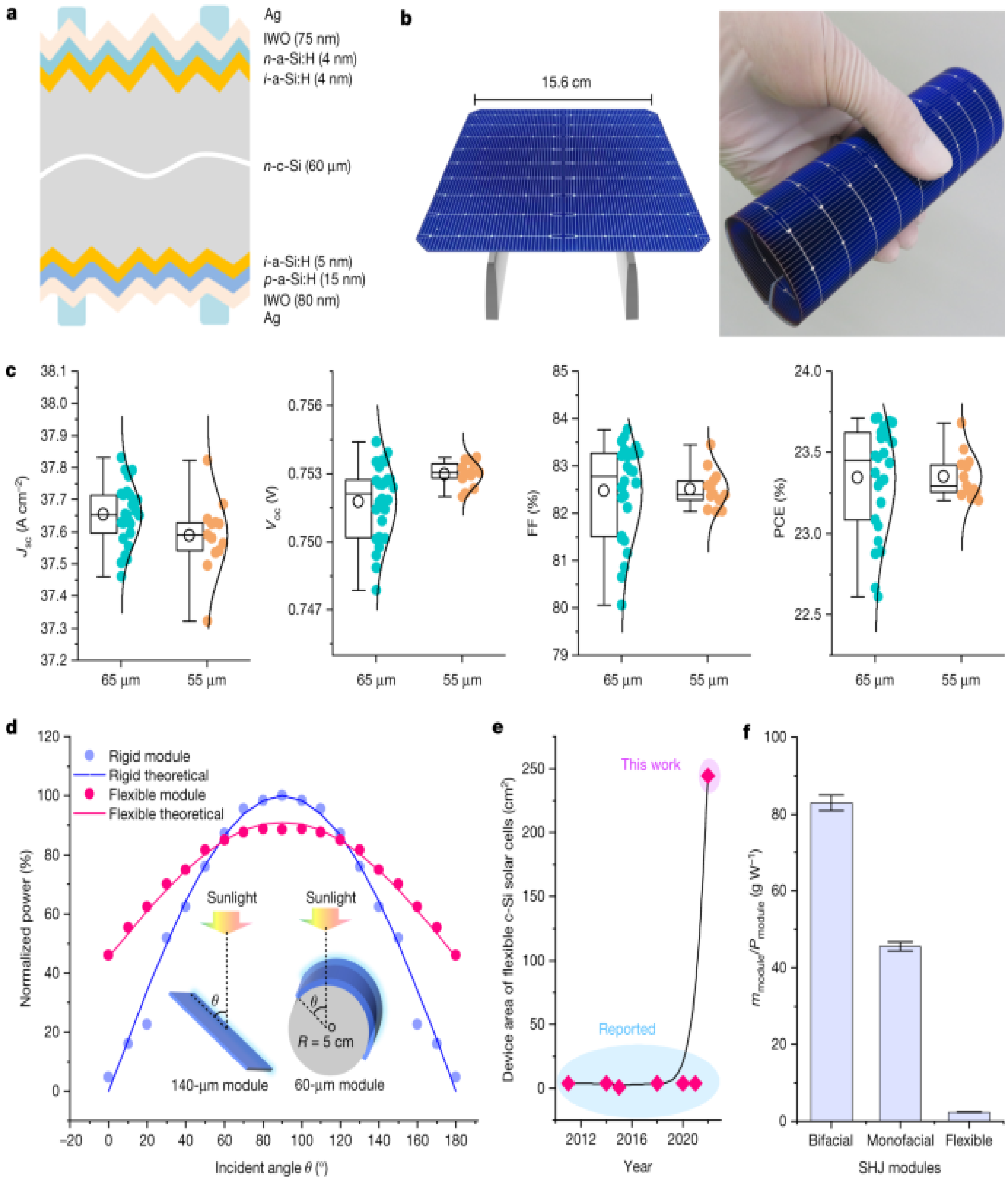
Solar cell (module) characterization

Next, we fabricated the foldable c-Si wafers into solar cells. The most widely used industrial silicon solar cells include passivated emitter and rear cells[18], tunnelling oxide passivated contact[19] solar cells and amorphous–crystalline silicon heterojunction[20] (SHJ) solar cells. As shown in Supplementary Fig. 6, unlike passivated emitter and rear cells and tunnelling

oxide passivated contact solar cells, which have asymmetric structure designs and are fired at a high temperature of 800 ± 20 °C, SHJ solar cells have a symmetric structural designs and are fired at a low temperature of 180 ± 5 °C. Therefore, SHJ technology is more suitable for manufacturing flexible solar cells because it is free from edge warping caused by inner stress during the firing process.

Figure 3a shows the architecture of fabricated SHJ solar cells; their edges allow them to be rolled to more than 360° (Fig. 3b). The photovoltaic performance of the 65-μm and 55-μm devices is shown in Fig. 3c. The short-circuit current density (J_{sc}), open-circuit voltage (V_{oc}), fill factor (FF) and PCE are 37.65 ± 0.09 mA cm^{-2}, 0.752 ± 0.002 V, 82.40 ± 0.99% and 23.31 ± 0.33%, respectively, for the 65-μm device. The corresponding J_{sc}, V_{oc}, FF and PCE values for the 55-μm device are 37.59 ± 0.11 mA cm^{-2}, 0.753 ± 0.001 V, 82.51 ± 0.39% and 23.35 ± 0.13%, respectively. These PCEs are higher than the value of 19.67 ± 0.34% for flexible SHJ solar cells fabricated using 60-μm quasiplanar wafers because of the higher J_{sc} values of 65-μm (37.65 mA cm^{-2}) and 55-μm (37.59 mA cm^{-2}) devices compared with that of the 65-μm wafer (31.45 mA cm^{-2}; Supplementary Fig. 7). After capping a 110-nm MgF_2 antireflective layer on the side exposed to sunlight, we submitted one flexible cell to an independent test centre and obtained a certified PCE of 24.50% for a 244.3-cm^2 wafer (Extended Data Fig. 6). Although this value was lower than that (25.83%) of a thick cell (Extended Data Fig. 7) because it was affected by the inferior light-harvesting ability of the thinner wafer[21], it was a remarkable PCE compared with that of the current flexible solar cells fabricated from other cost-effective materials. However, considering the V_{oc} of 750 mV for a 98-μm wafer and implied V_{oc} of about 760 mV for 40-μm wafers[22,23], the PCE in this work should be further improved through better surface passivation.

Fig. 3: Solar cell (module) performance.



a, Schematic of the architecture of the SHJ solar cells used in this study. IWO, tungsten-doped indium oxide. **b**, Photographs of a 15.6 cm × 15.6 cm flexible SHJ solar cell. **c**, J_{sc}, V_{oc}, FF and PCE of 65-µm and 55-µm SHJ solar cells. The top lines, bottom lines, lines in the box, circles and boxes represent maximum values, minimum values, median values, mean values and 25–75% distributions, respectively. **d**, Normalized power of two mini-

modules tested as a function of the incident angle of light θ: a rigid module assembled from a 140-μm SHJ cell and a flexible module assembled from a 60-μm SHJ cell. The latter was attached to a black cylinder with a radius of 5 cm. The theoretical power of the rigid module was given by $P(\theta) = \sin\theta$, whereas that of the flexible module was given by $P(\theta) = 0.455 \times [1 + \sin\theta]$. We collected experimental data from 0° to 90°; other data from 90° to 180° were symmetrically obtained by applying $P(\theta) = P(180° - \theta)$. **e**, Evolution of the device area of flexible c-Si solar cells. **f**, Mass-to-power ratio of bifacial, monofacial and flexible SHJ modules tested under standard conditions, where m_{module} and P_{module} are the mass and power of the modules.

Two mini modules were assembled to compare their performance: a rigid module encapsulating a 140-μm SHJ cell and a flexible module encapsulating a 60-μm SHJ cell. The latter was attached to a black cylinder with a radius of 5 cm. Their power was measured as a function of the incident angle of light (Fig. 3d). Although the flexible module showed a lower power at normal incidence (90°), its integrated electricity generation from 0° to 180° was 17% greater than that of the rigid module. Given that the 140-μm wafer accounted for approximately 50% of the device cost, the use of a 60-μm wafer reduced the production cost by approximately 29%. Overall, the flexible technology developed in this study reduced the levelized cost of energy by approximately 39% (23%) at the solar cell (module) level. Moreover, recent years have seen a rapid decrease in the diameter of diamond wire saws from about 80 μm to 40 μm, which can successfully cut 115 ± 5 μm wafers with a high product yield. The ability to produce thinner wafers with less kerf loss should contribute to reducing CO_2 emissions.

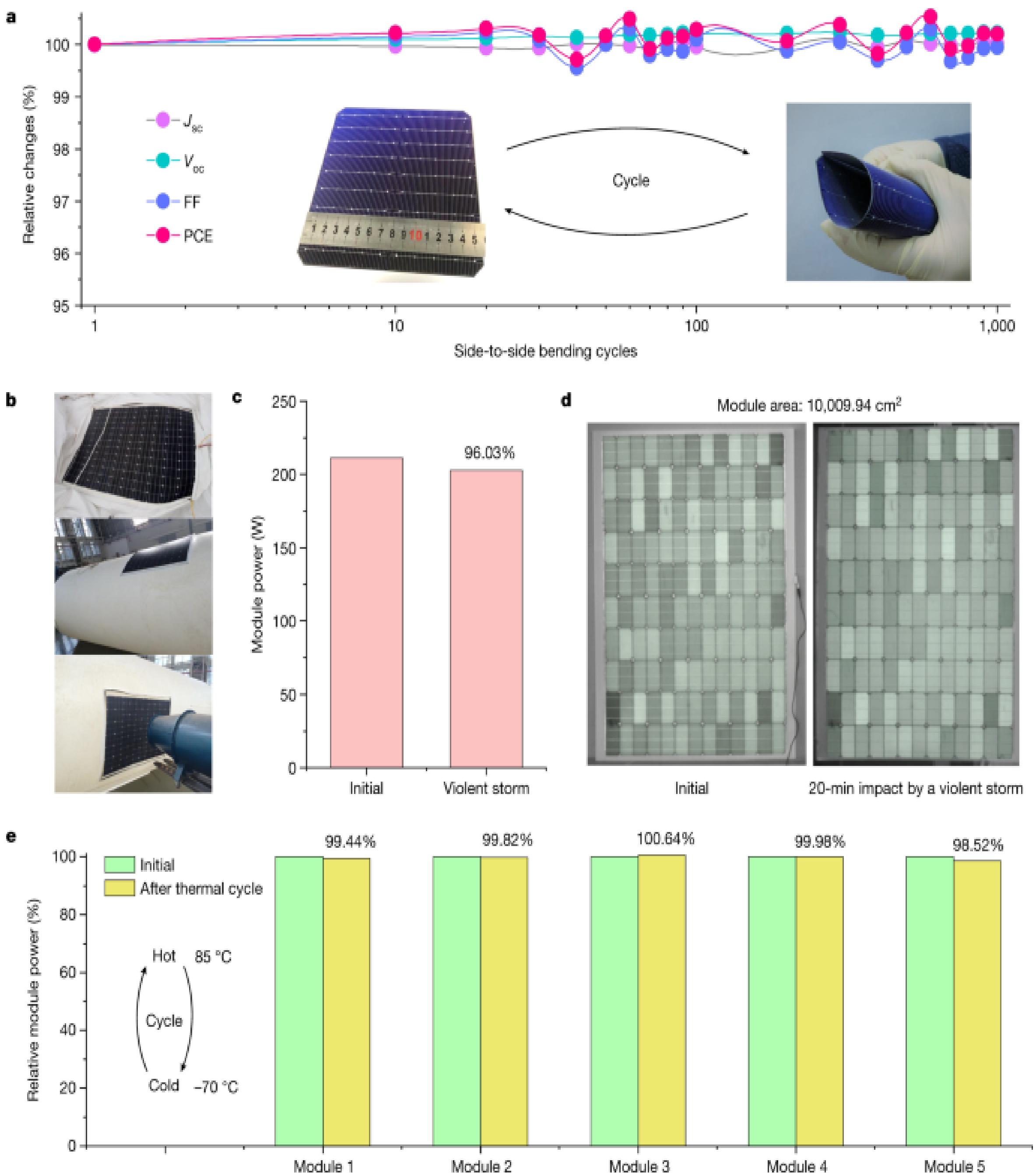
Extended Data Fig. 8 shows the rapid development of flexible solar cells during the past two decades[2,15,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38]. Our device highlights an advancement in the research field of flexible cells because most reported PCEs are below 20%. Specifically, the PCE of flexible c-Si cells has continuously increased over the past three years. In this study, we achieved remarkable increases in the device size and PCE from 4 cm^2 and 23.27% to 244.3 cm^2 and 24.5%, respectively (Fig. 3e and Extended Data Figs. 6 and 8). The realization of industrial-size, flexible c-Si solar cells indicates that the technology route demonstrated here is

compatible with standardized commercial production. At the module level, the flexible SHJ modules are free from heavy glasses and back sheets (Supplementary Figs. 8 and 9), which results in an extremely small mass-to-power ratio of 2.31 g W^{-1} that is much less than the values of 45.57 g W^{-1} and 82.93 g W^{-1} for standard monofacial and bifacial c-Si solar modules, respectively (Fig. 3f). The flexible SHJ modules demonstrated in this study may address the load-bearing issue encountered in the fast-growing research field of building-integrated photovoltaics and enable c-Si solar modules to be attached to building walls with either flat or curved surfaces.

Operating stability

Finally, we investigated the operating stability of the cell (module) under extreme conditions. The device exhibited a small R_b of approximately 8 mm (Supplementary Fig. 10). The J_{sc}, V_{oc}, FF and PCE of the flexible cell (Fig. 4a) retained 100% of their initial values after 1,000 side-to-side bending cycles. In each cycle, one edge was folded to touch the opposite edge; this bending was maintained for more than 10 s. The results for the bending cycles in the perpendicular direction are shown in Supplementary Fig. 11. These results were substantially different from that obtained for the flexible perovskite solar cell (Extended Data Fig. 8), in which the PCE decreased from 21% to 17% after the bending cycles. This decrease in PCE possibly originated from structural failure at grain boundaries in the polycrystalline perovskite film.

Fig. 4: Solar cell (module) stability.



a, Evolution of the performance of a flexible SHJ solar cell during a bending cycle. In each cycle, one edge was folded to touch the opposite edge; this bending was maintained for more than 10 s. **b**, A large (>10,000 cm^2) flexible SHJ solar module was attached to a soft gasbag. The pressure inside the gasbag was 94.7−830 Pa greater than the atmospheric pressure. Air was blown on the module by a fan to model a violent storm of 30 m s^{-1} for 20 min. **c**,**d**, The power of the module (**c**) and electroluminescence images (**d**)

before and after continuous air impact for 20 min to model a violent storm. **e**, The relative power of five flexible SHJ modules before and after thermal cycling between −70 °C and 85 °C for 120 h. In each cycle, the modules were maintained at −70 °C for 1 h and then at 85 °C for 1 h.

We assembled cells into a 10,009.94 cm^2 flexible module and attached this module to an inflated gasbag. Then, we used a powerful fan to model the effect of wind blowing at a speed of 30 m s^{-1} during a violent storm[39] (Beaufort number 11: 28.5−32.6 m s^{-1}; Fig. 4b and Supplementary Video 6). After continuous air impact for 20 min, the relative power loss was only 3.07% (Fig. 4c), which was consistent with the negligible changes in the electroluminescence images (Fig. 4d). This suggested that the flexible module can robustly operate under vibration conditions, which was also validated by the vibration cycles and free-falling cycles (Supplementary Figs. 12 and 13 and Supplementary Videos 7 and 8).

The lightweight nature of the flexible SHJ modules makes them suitable for charging near-space aerial vehicles[40], in which the temperature can reach as low as −70 °C at a height of 20−75 km. To model this, we cycled the flexible modules between −70 °C for 1 h and 85 °C for 1 h. After continuous temperature cycling for 120 h, the average relative power loss was only 0.32% (Fig. 4e), which shows that these modules can safely be operated in cold near-space conditions or on the South Pole or North Pole, as well as in deserts during hot summers.

Although the results obtained here for our cells are promising, they are not sufficient to guarantee consistent stability under real-life operating conditions in which stressors may occur simultaneously. Thus, more in situ tests need to be performed before large-scale production. So far, we have installed our flexible SHJ modules on near-space unmanned aerial vehicles (Extended Data Fig. 9) and on the South Pole (Supplementary Fig. 14). They can steadily supply electricity under extreme conditions and the output power is sometimes greater than the designed values, which probably results from the low temperature[41] and bifacial characteristics[42,43].

We showed that the mechanical performance of a material is not exclusively determined by its lattice structure at the atomic level; the symmetry at the mesoscale also plays an important part. At present, we can manufacture

more than 60,000 flexible SHJ cells daily with a fragmentation rate of less than 2% in our production line. This demonstrates an inexpensive strategy for commercial production of high-performance flexible c-Si solar cells. This might lead to a notable growth of the flexible cell market in the near future. Moreover, the concept demonstrated here applies, but is not limited, to solar cell fabrication; it should also be of interest to the community of researchers interested in other flexible electronics[44,45,46].

Methods

Stress simulation

The solid mechanics module in COMSOL Multiphysics (v.5.6) was used to simulate the stress of a two-dimensional silicon wafer with the length and thickness set to 1 cm and 60 μm, respectively. The Young's modulus, Poisson's ratio and mass density of the wafer were 130 GPa, 0.26 and 2.33 g cm^{-3}, respectively. The lower surface was textured with pyramids ranging from 5 μm to 8 μm in height and the initial angle between adjacent sharp pyramids was 71°. Three points around the midpoint of the top side of the wafer were fixed and bending forces of $F_b = 1.2$ mN were loaded on its two end points. The maximum von Mises stress was simulated as a function of the channel radius (R_p).

Atomistic simulation

Large-scale atomic/molecular massively parallel simulator (LAMMPS) package[47] was used to perform atomistic simulations of mode I loading on c-Si nanofilms with sharp and round channels between surface pyramids. The Tersoff potential[48] was used to describe the interatomic interaction between Si atoms. The simulated samples were 217.24 nm × 54.21 nm × 2.17 nm in size, containing approximately 1,150,000 Si atoms oriented along the [100], [010] and [001] directions with respect to the *x*, *y* and *z* axes, respectively. The R_p of the channels between the pyramids was increased from 0 to 15.81 nm. Periodic boundary conditions were imposed in the *y* and *z* directions of the simulation systems. Mode I loading was performed by uniformly stretching the simulation box with a strain rate of 5×10^8 s^{-1}.

Deformation processes coloured by the von Mises shear strain as well as stress−strain curves of the simulation samples with sharp and blunt notches were recorded as videos. Cracking of the blunted sample initiated at a higher loading strain of 17.3% compared with that of 9.3% for the untreated sample. Here, the simulation was qualitative because the R_p values were much smaller than those in the experimental conditions.

TEM characterization

An in situ bending test of a c-Si foil was conducted on an FEI Tecnai F30 TEM system using an electrical holder from PicoFemto. The c-Si foil was 6 μm × 12 μm × 70 nm in size, which was cut from the top surface of a wafer with sharp pyramids using a ThermoFisher Scios 2 FIB–SEM system, followed by deposition of a Pt film on the surface to protect the sharp pyramids. Then, the c-Si foil was welded onto a copper FIB holder with a diameter of 3 mm. A tungsten tip was used to contact the left side of the FIB c-Si foil; the movement of the foil was controlled by a piezo manipulator at a rate of about 0.01 nm s^{-1}, to apply a bending force on the edge of the c-Si foil with an estimated strain rate of 10^{-3} s^{-1}. For all bending processes, a 300-kV voltage with a weak electron beam was used in the TEM system to minimize the potential beam effects on the bending deformation. The real-time stress distribution was recorded by a charge-coupled device camera at a rate of 20 frames per second.

The fracture surfaces of two 60-μm wafers with sharp and round channels between pyramids were protected by a bilayer consisting of carbon and Pt films. In particular, a carbon film with a thickness of 100 nm was deposited by magnetron sputtering (ISC150 T Ion Sputter Coater) for nondestructive protection of the surface; then, the FIB-TEM foils were cut from the fracture surface using a ThermoFisher Scios 2 FIB–SEM system. STEM-HAADF observations were conducted at a depth of dozens of atoms from these fracture surfaces on an FEI Themis Z with a spherical aberration corrector for the illumination system.

Three-point bending test

Load–vertical displacement (*F*–*D*) curves of 4 cm × 2 cm × 140 μm textured c-Si wafers were obtained using a commercial Discovery DMA 850 instrument (Supplementary Fig. 15). The marginal regions of these textured wafers were blunted in 10 vol% HF:90 vol% HNO_3 solution for 0, 15 and 30 s.

GPA

The elastic strain distribution in the fractured c-Si wafers was mapped using GPA on the basis of the individual high-resolution STEM images. GPA, which was done on the basis of the formalism given in the literature[17] and implemented in the Gatan Digital Micrograph as a plug-in, was used to calculate the in-plane components of the symmetric strain tensor ε_{ij}. Strain maps were plotted with respect to an internal reference lattice based on g1 = (200) and g2 = (020) using Lorentzian masks with a diameter of 0.5 nm^{-1} (in reciprocal space). The maximum and minimum strains were set in the range of 5% to −5%.

SEM characterization

Top views, side views and fracture surfaces of c-Si wafers were observed using SEM (HITACHI, SU8020). Sharp channels between the pyramids of these wafers were blunted in 10 vol% HF:90 vol% HNO_3 solution for 0, 10, 20, 30, 40 and 90 s. The concentrations of HF and HNO_3 were 49% and 68%, respectively, diluted in water.

Ultraviolet–visible–infrared light characterization

The reflectivity of c-Si wafers from 300 to 1,200 nm was characterized using a UV−VIS−IR instrument (PerkinElmer Lambda 950).

Optical simulations

We used the electromagnetic wave module in COMSOL Multiphysics (v.5.6) to simulate the transmission, reflection and absorption spectra. A stack of a 10-nm a-Si:H layer and an 80-nm tungsten-doped indium oxide

layer was coated onto a 60-μm silicon slab. This structure was surrounded by air. Three silicon slabs were simulated; their surfaces were planar, pyramidal (height: 5 μm; pyramid angle: 71°) and rounded (radius: 2 μm). For the nonplanar silicon slabs, the thickness was the average distance between their boundaries. The upper and lower boundaries of the simulation region were set as Floquet boundary conditions. The wavelength and incident angle of light were distributed from 300 to 1,200 nm and from 0° to 80°, respectively. The plane wave entered from one side of the slab. The refractive index of air was 1, whereas the refractive indices of the other materials were analysed using ellipsometry. The transmittance and reflectance, defined as the ratio of the energy of the transmitted wave and reflected wave to that of the incident wave, respectively, were obtained by integrating the Poynting vector. The absorbance of the whole structure (silicon layer) was the ratio of the dissipation energy in the whole structure (silicon layer) to the energy of the incident wave.

Ultrahigh-speed video camera characterization

High-speed imaging of the cracking process of a 60-μm c-Si wafer with sharp pyramids was studied using a Phantom V2511 ultrafast CMOS video camera. It recorded up to 100,000 frames per second through a Leica Z16 APO long-distance microscope. The resolution was approximately 17.5 μm per pixel.

Solar cell fabrication

Czochralski *n*-type c-Si wafers were purchased from Sichuan Yongxiang. Their thickness and electrical resistivity were 160 μm and 0.3−2.1 Ω cm, respectively. The saw damage was removed in a 20.0-vol% alkaline water solution at 80 °C and the duration was varied to obtain different wafer thicknesses. Then, the wafers were textured in a 2.1-vol% alkali water solution at 80 °C for 10 min to form microscale pyramids on the surfaces. To fabricate flexible solar cells, the approximately 2-mm-wide marginal region of these 60-μm textured wafers was blunted in 10 vol% HF:90 vol% HNO_3 solution for 90 s at room temperature. All wafers were cleaned using a standard RCA process to remove organics and metal ions. Next, they underwent cleaning in 2.0% hydrofluoric acid water solution for 3 min to

etch the surface oxide. The creative thin c-Si technology developed previously has a great potential for flexible solar cells[49,50] because of sufficient utilization of the silicon material. Similar to the wet process, a dry method is also very efficient for improving the flexibility of the wafer (Supplementary Fig. 16). The marginal region of the wafer was blunted by a blending plasma (power, 120 W) of argon and fluorine ions for 30 min.

In a cluster plasma-enhanced chemical vapour deposition system (VHF-PECVD, IE Sunflower, OAK-DU-5; ULVAC CME-400), 5 nm *i*-a-Si:H and 15 nm *p*-a-Si:H, and 4 nm *i*-a-Si:H and 6 nm *n*-a-Si:H were deposited on the back and front sides of the wafers, respectively, in which the process temperatures were 200 ± 5 °C. The *i*-a-Si:H layers had a bilayer architecture; the first layer was grown using pure SiH_4, whereas the second layer was grown using diluted SiH_4 in H_2 with a flow ratio of 1:10. A 15-s H_2 plasma was used to improve the passivation quality at the interface of *i*-a-Si:H and *n*-c-Si. The power density, chamber pressure and gas flow ratio during deposition of the *n*-a-Si:H layer were 33 mW cm^{-2}, 80 Pa and $[PH_3]$:$[SiH_4]$:$[H_2]$ = 1.5:100:1,000, respectively. The *p*-a-Si:H layer also had a bilayer architecture, for which the deposition power density, chamber pressure and gas flow ratio were 20/20 mW cm^{-2}, 80/80 Pa and $[B_2H_6]$:$[SiH_4]$:$[H_2]$ = 1:100:100/2:100:400, respectively. Tungsten-doped indium oxide was deposited by reactive plasma deposition at 150 °C and the target was 1.0% tungsten dissolved in an indium oxide target. Electrode busbars and fingers were screen-printed on the surfaces of the devices using a low-temperature silver paste, followed by two-step annealing at 150 °C for 5 min and 185 °C for 30 min. On the sides of the certified SHJ solar cells that were exposed to sunlight, fine busbars were screen-printed and a 110-nm MgF_2 layer was deposited by electron-beam evaporation to improve the light-harvesting efficiency.

Solar cell (module) characterization

The current–voltage characteristics of SHJ solar cells (modules) were tested with a solar simulator (Halm IV, ceitsPV-CTL2) and the light intensity was calibrated using a National Renewable Energy Laboratory reference cell. A 60-μm flexible cell was independently tested by the National Institute of

Metrology in China. To compare the current density, a 140-μm brittle cell was independently tested by ISFH CalTeC in Germany. All devices were tested under a standard illumination of 100 mW cm^{-2} at 25 °C.

Bending cycle stability

The edge of a 60-μm flexible SHJ solar cell was folded to touch the opposite edge; this bending was maintained for more than 10 s. The bending speed was approximately 1,000 mm min^{-1}. The J_{sc}, V_{oc}, FF and PCE of this cell were tested with a solar simulator during 1,000 bending cycles under standard illumination of 100 mW cm^{-2} at 25 °C. The bending test was conducted in directions vertical and parallel to the direction of the busbars. We also monitored the sheet resistance of the 80-nm tungsten-doped indium oxide layer on a 60-μm quasiplanar c-Si substrate during 500 side-to-side bending cycles (Supplementary Fig. 17).

Vibration experiment

A 1,260 mm × 860 mm flexible SHJ module was installed on a large vibration platform, in which the module was supported by metal holders with a height of approximately 3 cm. The module vibrated in the *z* direction, expressed as $Z(t) = Z_0\sin(2\pi t/T)$, where the vibration amplitude $Z_0 = 5$ mm and the vibration period $T = 200$ ms. Electroluminescence images and the power of this flexible module before and after 18,000 vibration periods were obtained.

Free-falling test

We made a 5.4-kg, 520 mm × 520 mm module using our flexible SHJ solar cells, which was subjected 15 times to continuous free-falling from a height of approximately 500 mm. Its power was recorded before and after the free-falling cycles.

Thermal cycling stability

Thermal cycling was conducted between −70 °C and 85 °C for 120 h. In each cycle, the module was maintained at −70 °C for 1 h and at 85 °C for 1 h.

Violent storm stability

Flexible SHJ solar cells were encapsulated in a large (>10,000 cm^2) module, which was attached to a large soft gasbag inflated with air to support this flexible module. The pressure inside the gasbag was 94.7−830 Pa higher than the atmospheric pressure. A powerful fan was used to blow air on the module at a wind speed of 30 $m\ s^{-1}$ to model a violent storm (Beaufort number 11: 28.5−32.6 $m\ s^{-1}$). The power and electroluminescence images of this module before and after continuous impact by this air flow for 20 min were obtained.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All data generated or analysed during this study are included in the published Article and its Supplementary Information.

References

1. Crabb, R. L. & Treble, F. C. Thin silicon solar cells for large flexible arrays. *Nature* **213**, 1223–1224 (1967).

2. Hwang, I. et al. Effective photon management of non-surface-textured flexible thin crystalline silicon solar cells. *Cell Rep. Phys. Sci.* **1**, 100242 (2020).

3. Yoon, J. et al. Ultrathin silicon solar microcells for semitransparent, mechanically flexible and microconcentrator module designs. *Nat.*

Mater. **7**, 907–915 (2008).

4. Yoon, J. et al. Flexible concentrator photovoltaics based on microscale silicon solar cells embedded in luminescent waveguides. *Nat. Commun.* **2**, 343 (2011).

5. Ballif, C., Haug, F.-J., Boccard, M., Verlinden, P. J. & Hahn, G. Status and perspectives of crystalline silicon photovoltaics in research and industry. *Nat. Rev. Mater.* **7**, 597–616 (2022).

6. Söderström, T., Haug, F.-J., Terrazzoni-Daudrix, V. & Ballif, C. Optimization of amorphous silicon thin film solar cells for flexible photovoltaics. *J. Appl. Phys.* **103**, 114509 (2008).

7. Chirilă, A. et al. Potassium-induced surface modification of $Cu(In,Ga)Se_2$ thin films for high-efficiency solar cells. *Nat. Mater.* **12**, 1107–1111 (2013).

8. Metzger, W. K. et al. Exceeding 20% efficiency with in situ group V doping in polycrystalline CdTe solar cells. *Nat. Energy* **4**, 837–845 (2019).

9. Park, S. et al. Self-powered ultra-flexible electronics via nano-grating-patterned organic photovoltaics. *Nature* **561**, 516–521 (2018).

10. Blakers, A. W. & Armour, T. Flexible silicon solar cells. *Sol. Energy Mater. Sol. Cells* **93**, 1440–1443 (2009).

11. Park, H. et al. Improvement on surface texturing of single crystalline silicon for solar cells by saw-damage etching using an acidic solution. *Sol. Energy Mater. Sol. Cells* **93**, 1773–1778 (2009).

12. Sridharan, S., Bhat, N. & Bhat, K. N. Silicon surface texturing with a combination of potassium hydroxide and tetra-methyl ammonium hydroxide etching. *Appl. Phys. Lett.* **102**, 021604 (2013).

13. Saive, R. Light trapping in thin silicon solar cells: a review on fundamentals and technologies. *Prog. Photovolt.* **29**, 1125–1137 (2021).

14. A benchmark for visual identification of defective solar cells in electroluminescence imagery. elpv-dataset 1.0. GitHub. https://github.com/zae-bayern/elpv-dataset (2021).

15. Hwang, I., Um, H.-D., Kim, B.-S., Wober, M. & Seo, K. Flexible crystalline silicon radial junction photovoltaics with vertically aligned tapered microwires. *Energy Environ. Sci.* **11**, 641–647 (2018).

16. Pan, J., Ivanov, Y. P., Zhou, W. H., Li, Y. & Greer, A. L. Strain-hardening and suppression of shear-banding in rejuvenated bulk metallic glass. *Nature* **578**, 559–562 (2020).

17. Hÿtch, M. J., Snoeck, E. & Kilaas, R. Quantitative measurement of displacement and strain fields from HREM micrographs. *Ultramicroscopy* **74**, 131–146 (1998).

18. Blakers, A. Development of the PERC solar cell. *IEEE J. Photovolt.* **9**, 629–635 (2019).

19. Richter, A. et al. Design rules for high-efficiency both-sides-contacted silicon solar cells with balanced charge carrier transport and recombination losses. *Nat. Energy* **6**, 429–438 (2021).

20. Liu, W. et al. Light-induced activation of boron doping in hydrogenated amorphous silicon for over 25% efficiency silicon solar cells. *Nat. Energy* **7**, 427–437 (2022).

21. Dwivedi, N., Kumar, S., Bisht, A., Patel, K. & Sudhakar, S. Simulation approach for optimization of device structure and thickness of HIT solar cells to achieve similar to 27% efficiency. *Sol. Energy* **88**, 31–41 (2013).

22. Mikio, T. et al. 24.7% record efficiency HIT solar cell on thin silicon wafer. *IEEE J. Photovolt.* **4**, 96–99 (2014).

23. Balaji, P., Dauksher, W. J., Bowden, S. G. & Augusto, A. Development of 40 µm thin flexible silicon heterojunction solar cells. In *Proc. 2018 IEEE 7th World Conference on Photovoltaic Energy Conversion*. 2100–2103 (IEEE, 2018).

24. He, J. et al. Realization of 13.6% efficiency on 20 μm thick Si/organic hybrid heterojunction solar cells via advanced nanotexturing and surface recombination suppression. *ACS Nano* **9**, 6522–6531 (2015).

25. Chirilă, A. et al. Highly efficient Cu(In,Ga)Se_2 solar cells grown on flexible polymer films. *Nat. Mater.* **10**, 857–861 (2011).

26. Liu, D. & Kelly, T. L. Perovskite solar cells with a planar heterojunction structure prepared using room-temperature solution processing techniques. *Nat. Photon.* **8**, 133–138 (2014).

27. Lu, N. et al. 50-μm thick flexible dopant-free interdigitated-back-contact silicon heterojunction solar cells with front MoO_x coatings for efficient antireflection and passivation. *Opt. Express* **30**, 21309–21323 (2022).

28. Petermann, J. H. et al. 19%-efficient and 43 μm-thick crystalline Si solar cell from layer transfer using porous silicon. *Prog. Photovolt.* **20**, 1–5 (2012).

29. Qin, F. et al. Robust metal ion-chelated polymer interfacial layer for ultraflexible non-fullerene organic solar cells. *Nat. Commun.* **11**, 4508 (2020).

30. Ramanujam, J. et al. Flexible CIGS, CdTe and a-Si:H based thin film solar cells: a review. *Prog. Mater. Sci.* **110**, 100619 (2020).

31. Sai, H., Sato, Y., Oku, T. & Matsui, T. Very thin crystalline silicon cells: a way to improve the photovoltaic performance at elevated temperatures. *Prog. Photovolt.* **29**, 1093–1104 (2021).

32. Sai, H., Umishio, H. & Matsui, T. Very thin (56 μm) silicon heterojunction solar cells with an efficiency of 23.3% and an open-circuit voltage of 754 mV. *Sol. RRL* **5**, 21000634 (2021).

33. Sun, Y. et al. Flexible organic solar cells: progress and challenges. *Small Sci.* **1**, 2100001 (2021).

34. Tang, Q. et al. Investigation of optical and mechanical performance of inverted pyramid based ultrathin flexible *c*-Si solar cell for potential application on curved surface. *Appl. Surf. Sci.* **504**, 144588 (2020).

35. Wang, L. et al. Development of a 16.8% efficient 18-μm silicon solar cell on steel. *IEEE J. Photovolt.* **4**, 1397–1404 (2014).

36. Wu, C. et al. $FAPbI_3$ flexible solar cells with a record efficiency of 19.38% fabricated in air via ligand and additive synergetic process. *Adv. Funct. Mater.* **29**, 1902974 (2019).

37. Xie, H. et al. Recent progress of flexible perovskite solar cells. *Phys. Status Solidi Rapid Res. Lett.* **13**, 1800566 (2019).

38. Xu, Y. et al. Recent progress of electrode materials for flexible perovskite solar cells. *Nanomicro. Lett.* **14**, 117 (2022).

39. World Meteorological Organization. *Manual on Codes – International Codes, Volume I.1 Annex II to the WMO Technical Regulations Part A – Alphanumeric Codes* (WMO, 2019).

40. Young, M., Keith, S. & Pancotti, A. An overview of advanced concepts for near-space systems. In *Proc. 45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit* AIAA 2009-4805 (Curran Associates, 2010).

41. Xu, L. et al. Heat generation and mitigation in silicon solar cells and modules. *Joule* **5**, 631–645 (2021).

42. Han, C. et al. Towards bifacial silicon heterojunction solar cells with reduced TCO use. *Prog. Photovolt.* **30**, 750–762 (2022).

43. Guerrero-Lemus, R., Vega, R., Kim, T., Kimm, A. & Shephard, L. E. Bifacial solar photovoltaics – a technology review. *Renew. Sustain. Energy Rev.* **60**, 1533–1549 (2016).

44. Foster, M. A. et al. Silicon-chip-based ultrafast optical oscilloscope. *Nature* **456**, 81–84 (2008).

45. Yoon, J. et al. GaAs photovoltaics and optoelectronics using releasable multilayer epitaxial assemblies. *Nature* **465**, 329–333 (2010).

46. Chen, S. et al. Highly flexible and robust N-doped SiC nanoneedle field emitters. *NPG Asia Mater.* **7**, e157 (2015).

47. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).

48. Tersoff, J. Empirical interatomic potential for carbon, with applications to amorphous carbon. *Phys. Rev. Lett.* **61**, 2879–2882 (1988).

49. Niepelt, R. et al. High-quality exfoliated crystalline silicon foils for solar cell applications. *Energy Procedia* **55**, 570–577 (2014).

50. Hensen, J., Niepelt, R., Kajari-Schröder, S. & Brendel, R. Directional heating and cooling for controlled spalling. *IEEE J. Photovolt.* **5**, 195–201 (2015).

Acknowledgements

W.L. thanks J. Ren and J. Chen for their assistance with the free-falling experiment and three-point bending test, respectively. W.L. acknowledges financial support from the National Natural Science Foundation of China (grant no. 62004208) and Science and Technology Commission of Shanghai Municipality (grant no. 22ZR1473200). L.Z. acknowledges financial support from the National Natural Science Foundation of China (grant no. 62074153). B.D. acknowledges financial support from the National Natural Science Foundation of China (grant no. 12102021). A.H. acknowledges financial support from the Science and Technology Commission of Shanghai Municipality (grant no. 19DZ1207602). Y.F. acknowledges financial support from the National Natural Science Foundation of China (grant no. 62022085). D.L. acknowledges financial support from the Zhangjiang Laboratory. Z.D. acknowledges financial support from the National Natural Science Foundation of China (grant no. 51925208).

Author information

Author notes

1. These authors contributed equally: Wenzhu Liu, Yujing Liu, Ziqiang Yang, Changqing Xu

Authors and Affiliations

1. Research Center for New Energy Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

   Wenzhu Liu, Xiaodong Li, Shenglei Huang, Jianhua Shi, Junling Du, Anjun Han, Yuhao Yang, Kai Jiang, Zhenfei Li, Yinuo Zhou, Qiang Shi, Guangyuan Wang, Lan Guo, Liping Zhang, Fanying Meng & Zhengxin Liu

2. University of Chinese Academy of Sciences, Beijing, China

   Wenzhu Liu, Xiaodong Li, Kai Jiang & Liping Zhang

3. Institute of Metals, College of Material Science and Engineering, Changsha University of Science and Technology, Changsha, China

   Yujing Liu & Xiaochun Liu

4. Division of Physical Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

   Ziqiang Yang & Sigurdur Thoroddsen

5. Division of Computer, Electrical and Mathematical Science and Engineering, King Abdullah University of Science and Technology, Thuwal, Saudi Arabia

   Changqing Xu

6. School of Physical Science and Technology, ShanghaiTech University, Shanghai, China

Shenglei Huang & Lan Guo

7. Tongwei Solar Company, Chengdu, China

Jianhua Shi, Junling Du, Anjun Han, Shihu Lan, Haoxin Fu, Bin Fan, Guoqiang Xing, Yi Xie, Liping Zhang, Fanying Meng & Zhengxin Liu

8. Aerospace Information Research Institute, Chinese Academy of Sciences, Beijing, China

Guoning Xu, Yanchu Yang, Zhaojie Li & Rong Cai

9. Institute of Photovoltaics, Southwest Petroleum University, Chengdu, China

Jian Yu

10. UISEE Technologies, Shanghai, China

Jiajia Ling

11. Jiangsu Key Laboratory of Carbon-Based Functional Materials and Devices, Institute of Functional Nano and Soft Materials, Soochow University, Suzhou, China

Jun Peng

12. Institute of Solid Mechanics, Beihang University, Beijing, China

Liping Yu, Bin Ding & Yuan Gao

13. State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

Yanyan Fu

14. Key Laboratory of Wireless Sensor Networks and Communications of CAS, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

Wei He & Fengrong Li

15. School of Materials Science and Engineering, Jiangsu Collaborative Innovation Center of Photovoltaic Science and Engineering, Changzhou University, Changzhou, China

Xin Song

16. Paul-Drude-Institut für Festkörperelektronik, Leibniz Institut, Berlin, Germany

Jingxuan Kang

17. College of Energy, Soochow Institute for Energy and Materials Innovations, Soochow University, Suzhou, China

Xinbo Yang

18. The Interdisciplinary Research Center, Shanghai Advanced Research Institute, Chinese Academy of Sciences, Shanghai, China

Dongdong Li

19. Polar Research Institute of China, Shanghai, China

Zhechao Wang, Jie Li & Fuhai Wei

20. State Key Laboratory of Functional Materials for Informatics, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Sciences, Shanghai, China

Zengfeng Di

Contributions

W.L. conceived the idea, designed the overall experiments, led the project and wrote the paper. W.L., X. Li, S.H., J.S., J.D., A.H., Yuhao Yang, H.F., B.F., G. Xing,Y.X. and S.L. performed material and device optimization. X. Liu and Y.L. performed the TEM characterization and analysis. Z.Y. and S.T. performed the Phantom V2511 ultrafast-CMOS video camera characterization. C.X. performed the FEM simulations. G. Xu, Yanchu Yang, Zhaojie Li, R.C., Q.S. and W.L. performed the module stability tests. Z.W., J. Li and F.W. collected the data from the South Pole. B.D. and Y.G. performed the molecular dynamics simulations. Y.F., W.H., F.L., J.Y., Q.S., J. Ling, J.P., L.Y., B.D., K.J., Zhenfei Li, J. Li, X.S., Y.Z., G.W., L.G., J.K., D.L. and X.Y. contributed to material characterization and participated in data interpretation. Z.D. and Z. Liu discussed the results and co-wrote the paper. W.L., X. Liu, L.Z., F.M., Z.D. and Z. Liu supervised this project. All authors contributed to the discussion of the results and revision of the paper.

Corresponding authors

Correspondence to Wenzhu Liu, Xiaochun Liu, Liping Zhang, Fanying Meng, Zengfeng Di or Zhengxin Liu.

Ethics declarations

Competing interests

Shanghai Institute of Microsystem and Information Technology is in the process of applying for a patent application [202211090758.X] covering a method to fabricate flexible c-Si solar cells that lists W.L and Z. Liu as inventors. J.S., L.Z., A.H., J.D., S.L., H.F., B.F., G. Xing, Y.X., F.M. and Z. Liu. are employees of Tongwei Solar. J. Ling is an employee of UISEE Technologies. All other authors declare no competing interests.

Peer review

Peer review information

Nature thanks Bram Hoex, Robby Peibst and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Deformation and stress calculated by the solid mechanics module in COMSOL Multiphysics.

a, 2D simulation of a textured c-Si wafer whose length and thickness were set as 1 cm and 60 μm, respectively. Three points near the middle of the upper surface were fixed to exclude rigid-body displacement. Vertical forces F_b = 1.2 mN were loaded at the endpoints of the wafer. The bright colour indicates that most of the von Mises stress was concentrated in the sharp channels between pyramids. **b**, The maximum von Mises stress was simulated as a function of the channel radius R_p, which was tuned by a blunting treatment in an acid solution.

Extended Data Fig. 2 *In situ* bending test by TEM.

a, A stress holder for the *in situ* bending test during TEM. **b**, A tungsten tip controlled by the piezomanipulator was used to handle the left side of the FIB c-Si foil. **c**, Before loading the bending force, the bend contours were randomly distributed on the c-Si foil. **d**, After loading the bending force, these bend contours collected in the sharp channels between pyramids (white arrows), indicating that most stress concentrated in these sharp channels. The yellow squares implied that most stress shifted to nearby sharp channels after the bending force was loaded. Here, the dark fringes resulted from stress-induced lattice deformations.

Extended Data Fig. 3 Cracking process recorded by an ultrahigh-speed video camera.

Photographs of cracking of a 60 μm textured c-Si wafer on a time scale of 0−151 μs. The vertical arrows point to three emitted silicon particles, while the lateral arrows point to the endpoint of an evolving fracture. The circle highlights the site of initiation of the cracking.

Extended Data Fig. 4 Morphology of fracture surfaces.

a, Fracture surface of a 60 μm c-Si wafer with sharp channels between pyramids. **b**, Fracture surface of a 60 μm c-Si wafer whose channels between pyramids were blunted in 10 vol% HF/90 vol% HNO_3 solution for 30 s. The multiple cleavage sites in the fracture consumed more energy before the initiation of cracking.

Extended Data Fig. 5 Simulated fracture characteristics.

Together with the increase of R_p from 0 to 15.81 nm, the fractured path dually becomes more tortuous at a strain of 27.5%. Some jagged notches are also observed, in consistent with experimental results in Extended Data Fig. 4b.

Extended Data Fig. 6 Certificate report of a flexible SHJ solar cell.

Certificate of test results for a 60 μm flexible SHJ solar cell, whose short-circuit current density was 39.05 mA·cm^{-2} under AM1.5 illumination.

Extended Data Fig. 7 Certificate report of brittle SHJ solar cells.

a, Certificate of test results for a 140 μm brittle SHJ solar cell, whose short-circuit current density was 40.53 mA·cm^{-2} under AM1.5 illumination. **b**, Certificate of test results for a 140 μm brittle SHJ solar cell with Ag reflector

on the rear side, whose short-circuit current density was 40.95 mA·cm^{-2} under AM1.5 illumination.

Extended Data Fig. 8 PCE of flexible solar cells.

Emerging PCEs of flexible solar cells in the literature. Bending cycles decreased the PCE of the perovskite cell from 21% to 17%. For comparison, the certified PCE in this study of a 244.3 cm^2 c-Si wafer is also displayed. The dashed line indicates an efficiency boundary of 20%.

Extended Data Fig. 9 Application of flexible SHJ solar cells on a near-space unmanned aerial vehicle.

We installed lightweight flexible SHJ modules on a near-space unmanned aerial vehicle that safely flew at a height of 20 km thanks to the characteristics of the SHJ modules, which included light weight, flexibility and stability in cold environments.

Supplementary information

Supplementary Information

This file contains Supplementary Figs. 1–17 and Supplementary Table 1.

Reporting Summary

Supplementary Notes

This zipped folder contains Certificate Reports 1–3 and Vibrational Test Report. Descriptions of the four reports are also provided.

Supplementary Video 1

Shaking a flexible wafer. We can violently shake a flexible silicon wafer like a sheet of flexible paper.

Supplementary Video 2

Simulated cracking of a wafer with sharp pyramids. Atomistic simulations found the fracture initiated under a loading strain of 9.3%. The fracture surface was smooth.

Supplementary Video 3

Simulated cracking of a wafer with blunted pyramids (R_p = 5.27 nm). Atomistic simulations found the fracture initiated under a loading strain of 12.0%. The fracture surface was slightly rough.

Supplementary Video 4

Simulated cracking of a wafer with blunted pyramids (R_p = 10.54 nm). Atomistic simulations found the fracture initiated under a loading strain of 13.8%. The fracture surface became much rougher, in which jagged notches were observed along the complicated fracture path.

Supplementary Video 5

Simulated cracking of a wafer with blunted pyramids (R_p = 15.81 nm). Atomistic simulations found the fracture initiated under a loading strain of 17.3%. The fracture surface was the roughest, in which jagged notches were observed along the most tortuous fracture path.

Supplementary Video 6

Impact of a violent storm. The flexible SHJ module was blown by a violent storm.

Supplementary Video 7

Vibration cycles. The flexible SHJ module was vibrated in the vertical direction.

Supplementary Video 8

Free-falling cycles. The flexible SHJ module experienced free-falling cycles.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Liu, W., Liu, Y., Yang, Z. *et al.* Flexible solar cells based on foldable silicon wafers with blunted edges. *Nature* **617**, 717–723 (2023). https://doi.org/10.1038/s41586-023-05921-z

- Received: 24 August 2022
- Accepted: 06 March 2023
- Published: 24 May 2023
- Issue Date: 25 May 2023

- DOI: https://doi.org/10.1038/s41586-023-05921-z

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

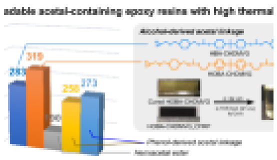
Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

Flexible solar cells made with crystalline silicon

Nature Research Briefing 24 May 2023

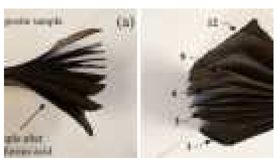
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05921-z



Novel degradable acetal-linkage-containing epoxy resins with high thermal stability: synthesis and application in carbon fiber-reinforced plastics

22 November 2021

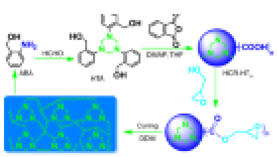
Mayo Kuroyanagi, Ayaka Yamaguchi, … Kazumasa Kawabe



High performance recycled CFRP composites based on reused carbon fabrics through sustainable mild solvolysis route

08 April 2022

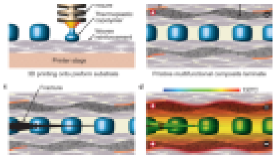
W. Ballout, N. Sallem-Idrissi, … P. Van Velthem



Recyclable thermoset hyperbranched polymers containing reversible hexahydro-s-triazine

09 December 2019

Zejun Xu, Yeyun Liang, … Menghe Miao



[Prolonged in situ self-healing in structural composites via thermo-reversible entanglement]

31 October 2022

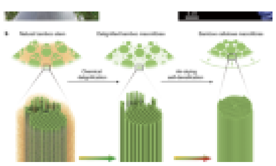
Alexander D. Snyder, Zachary J. Phillips, … Jason F. Patrick



[Effective reinforcements for thermoplastics based on carbon nanotubes of oil fly ash]

30 December 2019

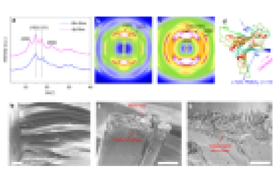
Numan Salah, Abdulrahman Muhammad Alfawzan, … Waleed Allafi



[Sustainable high-strength macrofibres extracted from natural bamboo]

16 December 2021

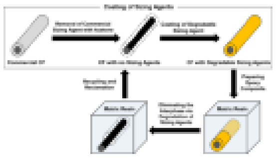
Zhihan Li, Chaoji Chen, … Liangbing Hu



[Integrating tough Antheraea pernyi silk and strong carbon fibres for impact-critical structural composites]

22 August 2019

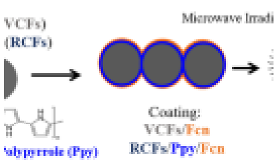
Kang Yang, Juan Guan, … Robert O. Ritchie



[Preparation of carbon fibers coated with epoxy sizing agents containing degradable acetal linkages and synthesis of carbon fiber-reinforced plastics (CFRPs) for chemical recycling]

10 May 2019

Ryosuke Aoki, Ayaka Yamaguchi, … Hirohumi Iyo



[Ultrafast carbon nanotubes growth on recycled carbon fibers and their evaluation on interfacial shear strength in reinforced composites]

02 March 2021

A. Salas, C. Medina, … M. F. Meléndrez

- Article
- Open Access
- Published: 26 April 2023

Catalytic disconnection of C–O bonds in epoxy resins and composites

- Alexander Ahrens ,
- Andreas Bonde,
- Hongwei Sun,
- Nina Kølln Wittig,
- Hans Christian D. Hammershøj,
- Gabriel Martins Ferreira Batista,
- Andreas Sommerfeldt,
- Simon Frølich,
- Henrik Birkedal &
- …
- Troels Skrydstrup

Show authors

Nature **volume 617**, pages 730–737 (2023)

- 17k Accesses

- 191 Altmetric

- Metrics details

Abstract

Fibre-reinforced epoxy composites are well established in regard to load-bearing applications in the aerospace, automotive and wind power industries, owing to their light weight and high durability. These composites are based on thermoset resins embedding glass or carbon fibres[1]. In lieu of viable recycling strategies, end-of-use composite-based structures such as wind turbine blades are commonly landfilled[1,2,3,4]. Because of the negative

environmental impact of plastic waste[5,6], the need for circular economies of plastics has become more pressing[7,8]. However, recycling thermoset plastics is no trivial matter[1,2,3,4]. Here we report a transition-metal-catalysed protocol for recovery of the polymer building block bisphenol A and intact fibres from epoxy composites. A Ru-catalysed, dehydrogenation/bond, cleavage/reduction cascade disconnects the C(alkyl)–O bonds of the most common linkages of the polymer. We showcase the application of this methodology to relevant unmodified amine-cured epoxy resins as well as commercial composites, including the shell of a wind turbine blade. Our results demonstrate that chemical recycling approaches for thermoset epoxy resins and composites are achievable.

Main

The vast amounts of end-of-use plastics and plastic-containing materials released into nature have resulted in a major environmental crisis[5,6], affecting ecosystems across the globe[9,10,11,12]. The necessity for the implementation of a circular economy of plastics and plastic-containing composites has become apparent for reducing the consumption of resources, as well as in limiting the introduction of waste into the environment[5]. In contrast to end-of-use thermoplastics, which can be melted and recast into new forms, the crosslinked polymer chains of thermoset plastics render these materials unsuitable for mechanical recycling. Circumventing processibility problems owing to lack of fusibility, chemical recycling deconstructs polymers into their original monomers or related base chemicals that can then re-enter established production chains yielding virgin polymeric materials. Enabling a circular economy in this fashion holds the opportunity of turning accumulating plastic waste into valuable resources[7]. Recently, the catalytic hydrogenation of thermoset polyurethane products for the recovery of anilines and polyols has been reported as a strategy realizing this principle[13,14]. By contrast, epoxy resins lack reactive carbonyl moieties, rendering selective disconnections of their chemical bonds more challenging. Lightweight, highly durable fibre-reinforced epoxy composites, which consist of glass or carbon fibres embedded in the polymer matrix, are high-performance materials crucial for the construction of automobiles, boats, aircraft and wind turbine blades[1]. Wind energy contributed to

approximately 6% of the global energy supply as of 2020, with projections forecasting significant growth in the near future[4]. In turn, 43 million metric tons of decommissioned wind turbine blades will have accumulated by 2050 (ref. [15]). At the same time, sustainable recycling technologies for such polymeric materials are almost nonexistent. Epoxy resins are not biodegradable and emit toxic gasses on incineration[16], ultimately leading to landfilling as the major pathway for their disposal. As of 2020, only around 1% of end-of-use composites were reused and this by means of shredding the material and using it as a filler substance in construction[1,2,3]. Because of its inefficiency and unsustainability, landfilling of wind turbine blades has been banned by several European countries, with more being expected to follow[4,17]. Hence, the pressing need for viable recycling strategies for epoxy resins and their composites is mounting[1,4].

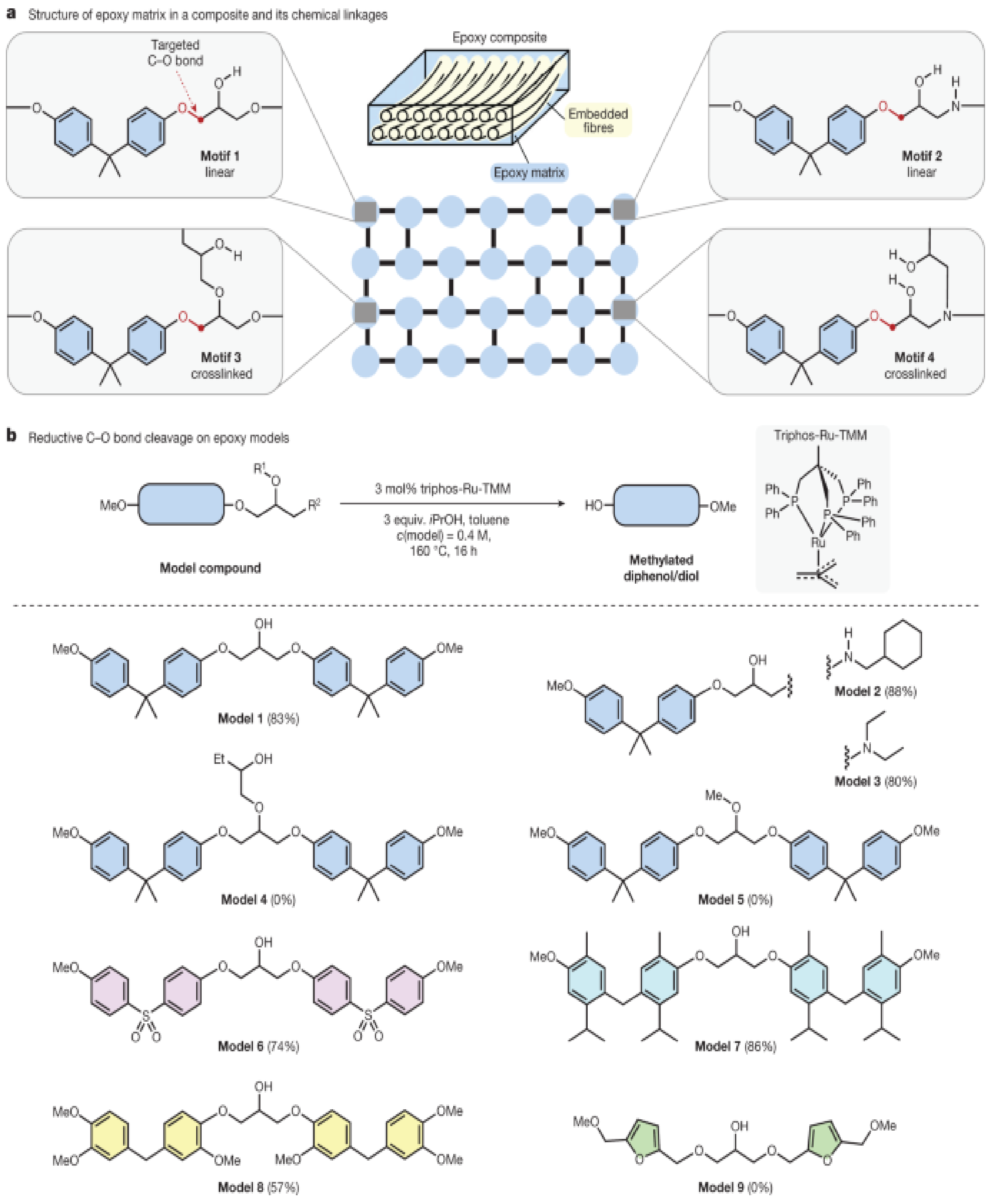
Methodologies investigated for recycling of polymer-based composites can be divided into two general approaches, both focused on recovery of fibres only. The first approach relies on destroying the polymer matrix by breaking chemical bonds unselectively, thereby releasing embedded fibres. Reported processes are based on harsh, energy-intense treatments, such as pyrolysis, which is impractical and results in damaged fibres[1,2,3]. Chemically destructive approaches yield fibres of higher quality[1] but require undesirable reagents such as hydrogen peroxide[18] or concentrated nitric acid[19]. The second, more elegant, approach is to design new epoxy resins containing 'molecular break points', which can be cleaved selectively under certain conditions[20,21]. Although the polymer matrix can be digested into soluble chain fragments, releasing the fibres, the recovered polymer fractions cannot be recast[22,23,24,25]. Furthermore, whereas the design of new resins could implement the reuse of fibres for future composite products, the legacy burden of epoxy materials produced up to the present day still remains, as well as those being produced now and in the near future using state-of-the-art resins.

We envisioned development of a chemical recycling approach for epoxy composites that aims at selective disconnection of the innate linkage motifs shared across epoxy resins, rather than cleaving artificially introduced breaking points or destroying the molecular complexity of the matrix. By targeting bonds formed during basic production steps of the resins, valuable

polymer building blocks could be recovered in addition to release of fibres from their polymeric embedment. Thereby, circularity for both epoxy resin and fibres would become achievable because virgin polymers may be produced from recovered base chemicals.

The petrochemical bisphenol A (BPA) represents an important building block for polymers, including epoxy resins[26]. For preparation of the latter, electrophilic epoxide moieties are attached to the BPA backbone via C(alkyl)–O single bonds. Difunctional epoxides can then be cured with multifunctional alkyl amines, yielding randomized three-dimensional (3D) polymer networks knitted together by C–O and C–N σ-bonds in different linkage motifs[27] (Fig. 1a). The potential leakage of BPA into the environment, and the ecological and human health risks associated with such events[28,29], have led to the investigation of potential substitutes[26]. However, with several million metrics tons of BPA-based materials in circulation to date, recovery of these materials and extraction of BPA in a controlled manner are highly desirable and serve to avoid potential leakages from landfilling sites. Biomass-based diphenol or diol motifs could be advantageous replacements for BPA, owing to their renewability[26]. Nonetheless, turning to biorenewable building blocks does not alleviate the requirement for a circular polymer production to minimize its environmental impact[30].

Fig. 1: Targeted C–O bonds in thermoset epoxy resins and catalytic deconstruction of related model compounds.



a, Schematic illustration of a crosslinked epoxy resin matrix and molecular structures of linkage motifs. Blue circles represent linkage sections while black lines represent linear polymer sections. The C–O bonds adjacent to BPA (red) are targeted to deconstruct the polymer matrix. **b**, Optimized reaction conditions applied to different model substrates, considering linkage

motifs and building blocks. Yields are given (in parentheses) for products isolated via column chromatography.

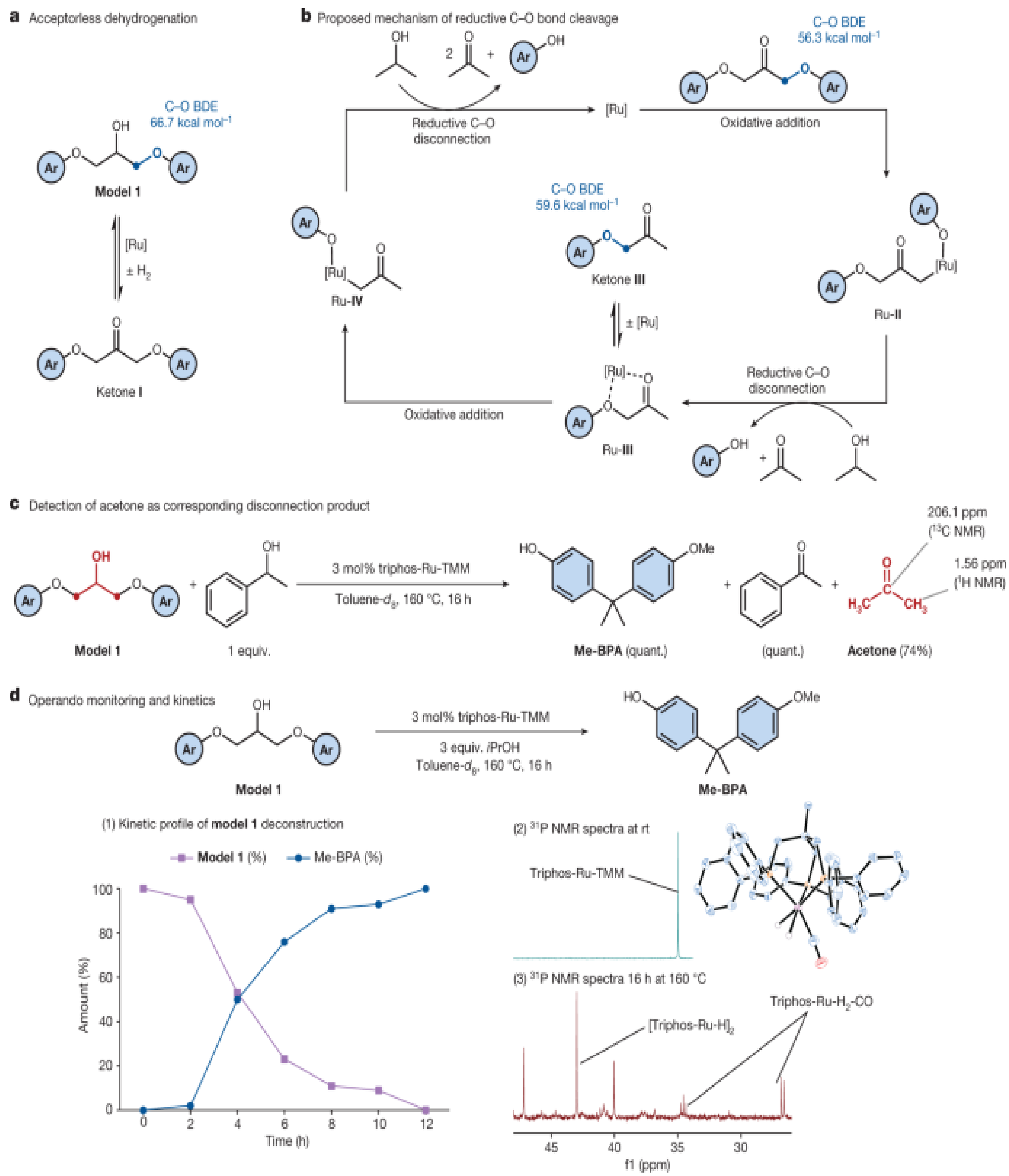
In pursuit of realizing the envisioned approach, we aimed at developing a transition-metal-catalysed method for disconnection of C(alkyl)–O single bonds adjacent to the BPA motif, which form during the reaction of BPA with epichlorohydrin. Carbon–oxygen single bonds have high bond dissociation energies[31,32] and their activation remains challenging. For the valorization of lignin[33], homogenous ruthenium catalysis has been reported to disconnect C(alkyl)–O and C–C single bonds[34,35,36]. Nonetheless, initial attempts to transfer these methodologies from lignin models to **model 1** (Fig. 1a), mimicking one of the most common motifs in epoxy resins, were unrewarding, leading to either low or no conversion (Supplementary Table 1). However, after a comprehensive screening of potential catalysts and reaction conditions (Supplementary Tables 2–6), an efficient protocol was identified. With triphos-Ru-TMM as the precatalyst and three equivalents of isopropanol in toluene at 160 °C, **model 1** was cleanly deconstructed providing methylated BPA (Me-BPA) in an isolated yield of 83% (Fig. 1b). No side products were detected from this process. Several other model compounds were tested, adapting the optimized deconstruction conditions. The efficient cleavage of **models 2** and **3**, yielding Me-BPA at 88% and 80%, respectively, shows that secondary and tertiary amines do not inhibit reactivity, an observation that is crucial when applying the methodology to relevant amine-cured epoxy resins. For substrate **model 4** mimicking the less common crosslinked motif 3, no conversion was observed. This linkage model requires the cleavage of three C–O bonds in total, whereby the first C–O bond cleavage is less feasible than for **model 1** owing to the more challenging liberation of an alkyl alcohol compared with a phenol. Lastly, **model 5**, with the central alcohol group capped as a methyl ether, proved unreactive to the deconstruction conditions.

Beyond linkage motifs, we considered model compounds containing bisphenol or diol scaffolds other than BPA (Fig. 1b). Bisphenol S (BPS) is a commercially relevant compound with a sulfonyl linker in its backbone, resulting in different electronic properties. The BPS-based **model 6** reacted cleanly under optimized conditions, allowing for a 74% isolated yield of methylated BPS. Next, three other model compounds containing potential

biorenewable substitutes for BPA[26] were investigated. We tested our disconnection approach on a representative selection of such candidates. **Model 7** is based on a bisphenol obtained from the terpenoid, carvacrol[37,38], that can be synthesized from limonene[39]. Under catalytic conditions, efficient and selective C–O bond cleavage was observed. Vanillin-derived compounds can be sourced from lignin, making them attractive candidates for replacement of petrochemical sources. An epoxy model based on *p*,*p'*-bisguaiacol F (**model 8**)[40] was subjected to catalytic deconstruction conditions and, after 16 h, methylated *p*,*p'*-bisguaiacol F was isolated at 57% yield. Although it can be speculated that the more sterically encumbered C–O bond is less accessible to the catalyst, the reaction is highly selective because only unreacted starting material was recovered apart from the cleavage product. Lastly, a bis(hydroxymethyl)furan-based model compound (**model 9**) was tested[41]. Furan-related building blocks are of interest because of their availability from cellulose[42]. No conversion was observed, showing that Ru-catalysed cleavage is selective for phenol-based linkages. With this scope of model substrates at hand, studies on catalytic activity in **model 1** were conducted. Even at a reduced catalyst loading of 0.38 mol%, quantitative conversion to Me-BPA was observed after 16 h. This corresponds to a turnover number (TON) of 533.

As supported by the inertness of **model 5** to the catalytic conditions, and in line with mechanisms proposed for cleavage on lignin β-O-4 linkage models[34,36], we postulate that C–O aryl bond breakage is preceded by dehydrogenation of the alcohol functionality in **model 1**, forming ketone **1** (Fig. 2a). The bond dissociation energies (BDEs) for the C–O ether bonds of **model 1** and the corresponding ketone (ketone **I**) were calculated. In alignment with theoretical studies on lignin linkages[31], the BDE of the latter is 10.7 kcal mol^{-1} lower (Fig. 2a,b), supporting the premise that dehydrogenation is necessary for initiation of C–O activation.

Fig. 2: Mechanistic considerations regarding Ru-catalysed C–O bond disconnection.



a, Ru-catalysed acceptorless dehydrogenation. **b**, Proposed catalytic cycle for disconnection of C–O bonds. BDEs were calculated with DFT at the (U)M06-2X/6-311++G(d,p) level of theory. **c**, Detection of acetone as the disconnection product. Yields were determined by ^{1}H NMR spectroscopy using 1,3,5-trimethoxybenzene as the internal standard. **d**, Kinetic profile of deconstruction of **model 1**. Operando monitoring experiment for C–O bond

disconnection on **model 1**. Molecular structure of triphos-Ru-H_2-CO in the crystal (CCDC 2219777).

From ketone **I**, we propose that C–O single bonds are susceptible to cleavage via an oxidative addition step involving a low-valent ruthenium complex eventually generating Ru-**II** (Fig. 2b). Through a dehydrogenation step with isopropanol, intermediate Ru-**II** is reduced with subsequent formation of species Ru-**III** and the liberation of a phenol. The low-valent ruthenium complex Ru-**III** undergoes a second oxidative addition step, and the reduction cascade ultimately leads to cleavage of the model substrate into acetone and phenol components. The kinetic profile of bond disconnection on **model 1** was investigated (Fig. 2d) showing an induction period of approximately 2 h, after which only a minute amount of Me-BPA could be detected. However, after 4 h, a 50% yield of Me-BPA was observed with quantitative conversion reached after 12 h. An induction period hints at catalyst activation preceding the catalytic cycle. At a reaction time of 4 h, traces of the monocleavage intermediate ketone **III** were detected. In that, otherwise, only the starting material and product were observed indicates that the consumption of intermediates is significantly faster than that of **model 1**. Density functional theory (DFT) calculations showed that the C–O BDE of ketone **I** is 3.3 kcal mol^{-1} lower than that of ketone **III** (Fig. 2b), implying that the more rapid consumption of the latter is not controlled by C–O bond strength. Possibly the second C–O bond activation is more rapid than dissociation of the intermediate from the catalyst. Ketone **III** was subjected to catalytic conditions in a one-to-one mixture with **model 1**. Both compounds were consumed yielding Me-BPA, which supports ketone **III** being an intermediate. To gather further support for the proposed C–O bond-cleaving mechanism, efforts were undertaken to detect acetone as the corresponding disconnection product formed from the central linkage motif of **model 1** (Fig. 2c). Because the dehydrogenation of isopropanol yields acetone, the hydrogen source was exchanged for 1-phenylethanol. After 16 h of reaction time, both a quantitative formation of acetophenone from 1-phenylethanol and 74% yield of acetone from the linkage motif were observed.

The deconstruction of **model 1** using 3 mol% triphos-Ru-TMM in toluene at 160 °C in the absence of isopropanol yielded only 4% of Me-BPA after 16 h

(Supplementary Table 1). If isopropanol serves only as an additional hydrogen source, approximately 50% of the C–O bonds should have been cleaved. To investigate the role of isopropanol, operando monitoring experiments were conducted (Fig. 2d) using ^{1}H and ^{31}P nuclear magnetic resonance (NMR) spectroscopy. The experiment, in the absence of isopropanol, was run in a J. Young NMR tube. After 16 h, no conversion of the starting material was observed. Likewise, triphos-Ru-TMM had not been consumed and no other ruthenium species could be detected. The reaction was repeated with 3 equiv. of isopropanol present. Here, close to quantitative conversion of the starting material to Me-BPA was observed after 16 h. In addition, the starting ruthenium complex was consumed and the formation of new signals corresponding to ruthenium species was detected in both the ^{31}P NMR spectra and the hydride region of the ^{1}H NMR spectra. The addition of another equivalent of **model 1** to the reaction mixture containing the new ruthenium species and running it again led to a clean deconstruction of the fresh substrate into methylated BPA.

A phosphine singlet peak at 47.3 ppm could not be identified. A hydride-bridged binuclear ruthenium(I) complex[43] was detected and tested as a potential precatalyst on **model 1**. With no conversion being observed, the formation of this species can be considered a deactivation pathway. Lastly, a triplet peak and doublet phosphine peak could be linked to a hydride signal using an ^{1}H, ^{31}P heteronuclear single quantum coherence NMR experiment. Single crystals suitable for X-ray crystallography were obtained, allowing the identification of this species as triphos-Ru-H_2-CO. We propose that the CO originates from acetone, which presents a rare case of decarbonylation of a secondary alcohol[44]. The carbonyl complex was tested as precatalyst, but proved inactive.

With the identification of a suitable catalytic system for the deconstruction of epoxy models, we moved on to a polymeric resin (Fig. 3). Airstone 760E/766H is a thermoset epoxy system adapted to the construction of wind turbine blades, consisting of four monomers and containing approximately 43 wt% of BPA after curing. Initially a ‘dogbone’ of the cured resin was ground into powder, suspended in toluene-containing isopropanol and stirred at 160 °C in the absence of catalyst (entry 1). After 24 h, no compounds liberated from the resin could be detected. By contrast, when 6 wt% of

catalyst was added, a 56% yield of BPA was isolated (entry 2). Both gas chromatography–mass spectrometry and ^{1}H NMR spectroscopic analysis of the BPA sample did not show any detectable impurities, supporting the potential for its reuse. Furthermore, a highly polar rest fraction was recovered, the analysis of which by matrix-assisted laser desorption/ionization time-of-light mass spectrometry, ^{1}H NMR and infrared spectroscopy showed a complex mixture of oligomers containing alkyl ethers, amines and small amounts of aromatics. In total, 81 wt% of the cured resin was recovered as soluble organic material. The influence of particle size on the efficiency of the deconstruction suggests that catalysis is limited to the surface area of the resin particles suspended in solution (Supplementary Table 7). An attempt to decrease catalyst loading by 50% to 3 wt% (entry 3) reduced the amount of recovered BPA to a 34% yield. However, extending the reaction time to 4 days (entry 4) increased the amount of recovered BPA to 81%, showing that the catalyst was not deactivated after 24 h. Another 4-day experiment using 1 wt% of catalyst provided a 42% yield of BPA corresponding to a TON of 105 (entry 5).

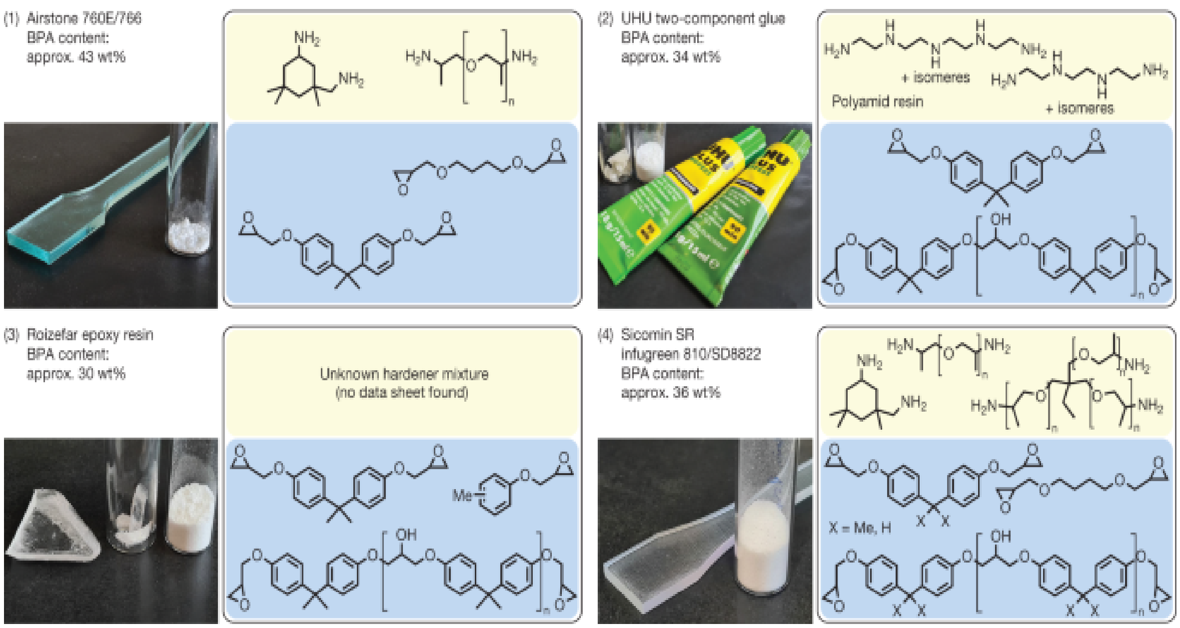
Fig. 3: Catalytic deconstruction of epoxy resins.

Epoxy resin ⟶ BPA + Remaining fraction

Reaction conditions:
6 wt% triphos-Ru-TMM
8 vol% *i*PrOH in toluene
c(epoxy) = 92 mg ml^{-1}, 160 °C, 24 h

Entry		Recovered BPA , % (mg)	Remaining fraction (mg)	Total mass (%)
1	(a) Airstone, no catalyst	0 (0)	–	–
2	(a) Airstone	56 (24.3)	60.1	81
3	(a) Airstone, 3 wt% catalyst	34 (14.6)	29.2	44
4	(a) Airstone, 3 wt% catalyst, 4 d	81 (35.0)	49.4	84
5	(a) Airstone, 1 wt% catalyst, 4 d	53 (22.6)	63.7	86
6	(b) UHU two-component glue	38 (12.9)	20.8	34
7[a]	(c) Roizefar epoxy resin	50 (15.0)	32.8	57
8	(d) Sicomin SR infugreen	54 (19.5)	48.3	68

[a] 9 wt% cresols were isolated together with BPA.



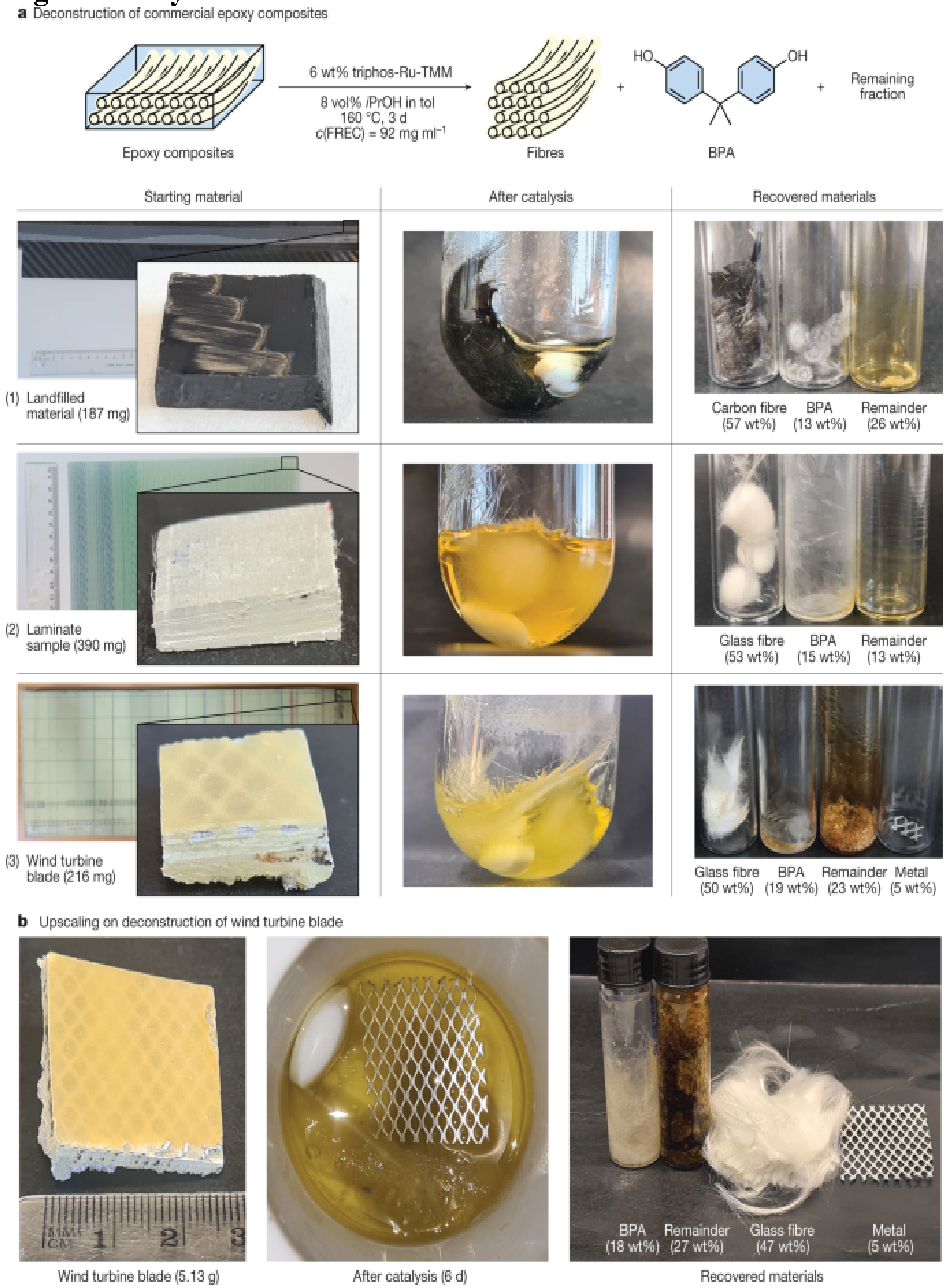
Scope of epoxy resins deconstructed using catalytic conditions. Experiments were set up under an argon atmosphere. Yields were determined after isolation of products via column chromatography.

Three additional commercial epoxy resins were tested. A two-component adhesive (entry 6), containing a complex mixture of multifunctional hardeners resulting in a higher degree of crosslinking, proved to be more challenging to deconstruct. Nonetheless, a 38% yield of BPA was recovered after 24 h. A clear-cast epoxy system marketed for handicrafts (entry 7) was also disassembled yielding 50% of BPA. Furthermore, this resin was found to contain a cresol-based epoxy component because two isomers of cresol were recovered at a yield of 9 wt% together with BPA. Next, a partially biomass-based infusion system for maritime engineering was tested (entry 8). This system comprises a complex mixture of multifunctional hardeners and contains small amounts of bisphenol F (BPF) epoxides. Although 54%

of the BPA was recovered from this sample, the separated BPF remained elusive after deconstruction possibly because of the modest amounts of BPF epoxides applied in this specific mixture. Last, Lightstone 3100E/3102H, an anhydride-curing system developed for pultrusion applications, was subjected to catalysis. However, for this sample no BPA could be isolated. Anhydride-curing results in the formation of linkage motifs different from those derived from amine curing, because the alcohol moieties of the epoxy fraction are acylated[27]. This functionalization blocks dehydrogenation and subsequently C–O bond cleavage analogously to **model 5**.

With a general method available for molecular disassembly of amine-cured epoxy resins, we turned to investigate the suitability of this protocol for the deconstruction of fibre-reinforced epoxy composites, which, apart from the polymer matrices, contain a high weight percentage of fibres. For recovery of the latter, powderizing epoxy composites are counterproductive. We recognized that composite materials may be more accessible to the solvent because of the interface area between fibres and polymer. Our work commenced with a carbon fibre-based composite ((1) in Fig. 4a), procured from landfill. With no previous treatment apart from cutting to size, a cube of material was submerged in the solvent mixture, 6 wt% of catalyst was added and the mixture stirred at 160 °C. After 3 days the composite had visibly separated into loose fibres. The reaction mixture was decanted; after washing, 57 wt% of carbon fibres was recovered and, from the solution, 13 wt% of BPA was isolated. Although we could not estimate the total amount of BPA in the material, the complete solubilization of the polymer matrix and mass balance point towards an efficient recovery. Next, a commercial product sample of a glass fibre-based laminate ((2) in Fig. 4a)) was treated in the same manner. Here too visible separation into loose fibres was observed over the course of 3 days. This composite yielded 53 wt% of glass fibres and 15 wt% of BPA, demonstrating that the method does not differentiate between glass and carbon fibre-based composites. Last, with these promising results at hand, a piece of the outer shell of a state-of-the-art decommissioned wind turbine blade ((3) in Fig. 4a) was tested. This commercial composite sample was cleanly disassembled, yielding 50 wt% of glass fibres and 19 wt% of BPA. Additionally a piece of metal grid, incorporated in the blade as part of the lightning protection system, was separated from the structure.

Fig. 4: Recovery of BPA and fibres from commercial epoxy composites using Ru catalysis.

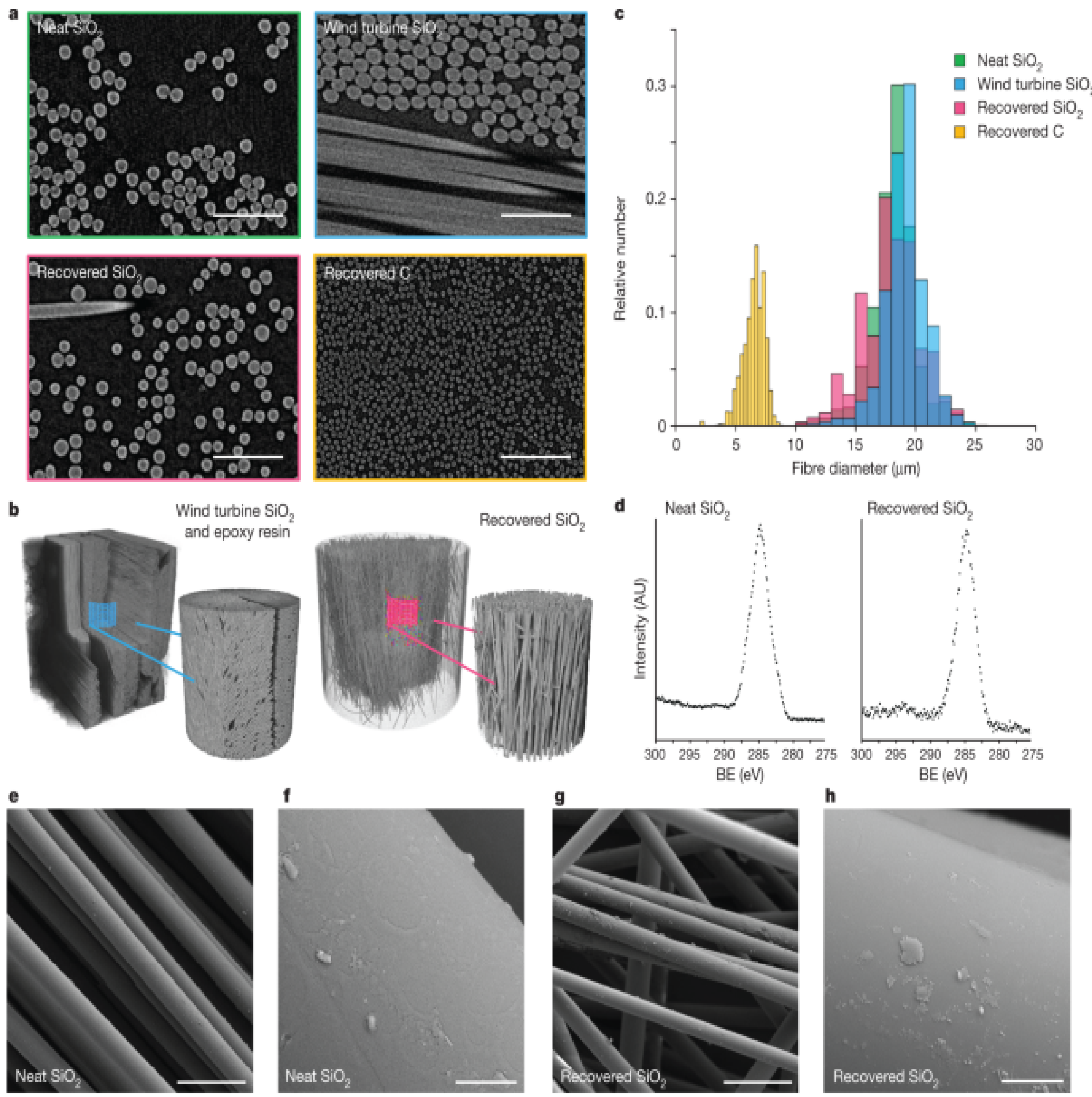


a, Scope of the composite samples subjected to catalysis. Composite pieces 1, 2 and 3 were 1.0–1.5 cm in both length and width. **b**, Upscaling of deconstruction conditions on wind turbine blade.

Finally, the possibility of scaling up the catalytic protocol to larger pieces of composite material was examined (Fig. 4b). A 5.13-g plate of decommissioned wind turbine blade was placed in a 300-ml autoclave and subjected to catalytic conditions. After 6 days of reaction time the matrix had been fully disassembled, allowing the isolation of 4.0 mmol of pure BPA and the recovery of 2.4 g of pristine glass fibres. Furthermore, a sheet of metal grid was also recovered.

To evaluate the quality of the recovered fibres in comparison with neat fibres, X-ray micro-computed tomography (µ-CT), X-ray photoelectron spectroscopy (XPS) and scanning electron microscopy (SEM) were used (Fig. 5). With µ-CT, grey-level variations in the images reflect varying material density. Glass fibres appeared lighter than the epoxy resin, which, in turn, appeared slightly lighter than air (Fig. 5a). Figure 5b shows 3D renderings of fibres where air has been made transparent. Epoxy resin was observed only in the nontreated piece of the wind turbine blade, and the images visually corroborate that the high quality of the recovered fibres was preserved. Diameters of the glass fibres were quantified and found to be similar to those of neat glass fibres (18 ± 2 µm), and for glass fibres both embedded within (19 ± 2 µm) and recovered from (18 ± 2 µm) a piece of decommissioned wind turbine blade. The carbon fibres salvaged from landfilled material (6.5 ± 0.9 µm) were substantially smaller (Fig. 5c).

Fig. 5: Characterization of commercial fibre-reinforced composites and glass fibres.



a, X-ray μ-CT with virtual slices through reconstructed image stacks showing fibre cross-sections. Scale bars, 100 μm. **b**, 3D renderings of reconstructed image stacks showing fibre organization; grey levels corresponding to air have been rendered transparent; for scale refer to the two-dimensional slices in **a**. **c**, Histograms of fibre diameter obtained by analysis of X-ray μ-CT data. **d**, XPS C 1s high-resolution spectra of neat and recovered fibres. **e**–**h**, SEM images of neat (**e**,**f**) and recovered fibres (**g**,**h**). Scale bars, 50 μm (**e**,**g**); 2 μm (**f**,**h**). AU, arbitrary units; BE, binding energy.

X-ray photoelectron spectroscopy was used to test whether the epoxy resin had been completely removed from the fibres in the deconstruction process

(Supplementary Table 9). The atomic concentrations of Si, Ca and Al relative to C were higher in the recovered than in the neat fibres. The higher relative C content of the latter originates from the priming layer used to coat glass fibres, which was partially removed during catalysis. No residual polymer was detected, and this was further corroborated by the high-resolution C 1s spectra of both neat and recovered fibres (Fig. 5d), for which the π-π* type shake-up peaks typically detected for C in aromatic compounds (around 291–292 eV) were absent. SEM images of fibres show the imprint of this coating on neat fibres (Fig. 5e,f and Supplementary Fig. 16a–d,i), whereas the surface of the recovered fibres is smooth (Fig. 4g,h and Supplementary Fig. 16e–h,j). Finally, preliminary tensile strength studies on fibres recovered from the wind turbine blade, with neat fibres as the reference point, showed comparable mechanical strength (Supplementary Table 10).

For those components recovered from end-of-use composites, perspectives for circularity can be considered. The high purity of the BPA recovered allows, in principle, its reuse in established production chains for epoxy resins, polycarbonates or polyesters, replacing virgin BPA produced from naphtha feedstock. The remaining fraction, consisting of various oligomers, unfortunately cannot be used as chemical building blocks. Nonetheless, valorization strategies beyond energy recovery can be envisioned. For example, pyrolysis has been demonstrated to process mixed plastic wastes (including nitrogen-containing polymers) into, for example, naphtha equivalents or syngas[45,46]. As such, this remaining fraction may find uses as a carbon feedstock source for the chemical industry. Last, in regard to the glass and carbon fibres recovered at high quality, several reuse approaches have been reported. These include the use of recovered fibres to construct new composite materials after a chemical sizing or resizing process[47,48]. With these considerations in mind, our catalytic process can be considered as a proof-of-concept demonstration that a circular economy may well be achievable for these valuable and relevant materials.

Methods

Descriptions of the methods used are provided in the Supplementary Information.

Data availability

Crystallographic data are available free of charge from the Cambridge Crystallographic Data Centre under no. CCDC 2219777. The authors declare that all other data supporting the findings of this study are available within the paper and its Supplementary Information files.

References

1. Navarro, C. A. et al. A structural chemistry look at composites recycling. *Mater. Horiz.* **7**, 2479–2486 (2020).

2. Oliveux, G., Dandy, L. O. & Leeke, G. A. Current status of recycling of fibre reinforced polymers: review of technologies, reuse and resulting properties. *Prog. Mater Sci.* **72**, 61–99 (2015).

3. Witik, R. A., Teuscher, R., Michaud, V., Ludwig, C. & Månson, J.-A. E. Carbon fibre reinforced composite waste: an environmental assessment of recycling, energy recovery and landfilling. *Compos. Part A Appl. Sci. Manuf.* **49**, 89–99 (2013).

4. Jacoby, M. Recycling wind turbine blades. *C&EN Global Enterprise* **100**, 26–30 (2022).

5. Kakadellis, S. & Rosetto, G. Achieving a circular bioeconomy for plastics. *Science* **373**, 49–50 (2021).

6. Geyer, R., Jambeck, J. R. & Law, K. L. Production, use, and fate of all plastics ever made. *Sci. Adv.* **3**, e1700782 (2017).

7. Vollmer, I. et al. Beyond mechanical recycling: giving new life to plastic waste. *Angew. Chem. Int. Ed. Engl.* **59**, 15402–15423 (2020).

8. Plastics Europe. Plastics – the Facts 2019 https://plasticseurope.org/knowledge-hub/plastics-the-facts-2019/ (2019).

9. Law, K. L. Plastics in the marine environment. *Annu. Rev. Mar. Sci.* **9**, 205–229 (2017).

10. Zhang, Y. et al. Atmospheric microplastics: a review on current status and perspectives. *Earth Sci. Rev.* **203**, 103118 (2020).

11. Hale, R. C., Seeley, M. E., La Guardia, M. J., Mai, L. & Zeng, E. Y. A global perspective on microplastics. *J. Geophys. Res. Oceans* **125**, e2018JC014719 (2020).

12. Stubbins, A., Law, K. L., Muñoz, S. E., Bianchi, T. S. & Zhu, L. Plastics in the Earth system. *Science* **373**, 51–55 (2021).

13. Gausas, L. et al. Catalytic hydrogenation of polyurethanes to base chemicals: from model systems to commercial and end-of-life polyurethane materials. *JACS Au* **1**, 517–524 (2021).

14. Zhou, W. et al. Depolymerization of technical-grade polyamide 66 and polyurethane materials through hydrogenation. *ChemSusChem* **14**, 4176–4180 (2021).

15. Liu, P. & Barlow, C. Y. Wind turbine blade waste in 2050. *Waste Manage.* **62**, 229–240 (2017).

16. Mulcahy, K. R., Kilpatrick, A. F. R., Harper, G. D. J., Walton, A. & Abbott, A. P. Debondable adhesives and their use in recycling. *Green Chem.* **24**, 36–61 (2022).

17. d'Ambrières, W. Plastics recycling worldwide: current overview and desirable changes. *Field Actions Sci. Re*p. **19**, 12–21 (2019).

18. Li, J. et al. A promising strategy for chemical recycling of carbon fiber/thermoset composites: self-accelerating decomposition in a mild oxidative system. *Green Chem.* **14**, 3260–3263 (2012).

19. Dang, W., Kubouchi, M., Yamamoto, S., Sembokuya, H. & Tsuda, K. An approach to chemical recycling of epoxy resin cured with amine using nitric acid. *Polymer* **43**, 2953–2958 (2002).

20. Liu, Y. et al. Closed-loop chemical recycling of thermosetting polymers and their applications: a review. *Green Chem.* **24**, 5691–5708 (2022).

21. Post, W., Susa, A., Blaauw, R., Molenveld, K. & Knoop, R. J. I. A review on the potential and limitations of recyclable thermosets for structural applications. *Polym. Rev.* **60**, 359–388 (2020).

22. Okajima, I., Hiramatsu, M., Shimamura, Y., Awaya, T. & Sako, T. Chemical recycling of carbon fiber reinforced plastic using supercritical methanol. *J. Supercrit. Fluids* **91**, 68–76 (2014).

23. Liu, T. et al. Selective cleavage of ester linkages of anhydride-cured epoxy using a benign method and reuse of the decomposed polymer in new epoxy preparation. *Green Chem.* **19**, 4364–4372 (2017).

24. Liu, T. et al. Mild chemical recycling of aerospace fiber/epoxy composite wastes and utilization of the decomposed resin. *Polym. Degrad. Stab.* **139**, 20–27 (2017).

25. Ma, Y., Navarro, C. A., Williams, T. J. & Nutt, S. R. Recovery and reuse of acid digested amine/epoxy-based composite matrices. *Polym. Degrad. Stab.* **175**, 109125 (2020).

26. Liguori, F., Moreno-Marrodan, C. & Barbaro, P. Biomass-derived chemical substitutes for bisphenol A: recent advancements in catalytic synthesis. *Chem. Soc. Rev.* **49**, 6329–6363 (2020).

27. Pham, H. Q. & Marks, M. J. in *Ullmann's Encyclopedia of Industrial Chemistry* (eds, Ley, C. et al.) https://doi.org/10.1002/14356007.a09_547.pub2 (Wiley-VCH, 2005).

28. Liu, J. et al. Occurrence, toxicity and ecological risk of bisphenol A analogues in aquatic environment – a review. *Ecotoxicol. Environ. Saf.* **208**, 111481 (2021).

29. Preethi, S. et al. Toxicity of bisphenol A on humans: a review. *Int. Lett. Nat. Sci.* **27**, 32–46 (2014).

30. Hong, M. & Chen, E. Y. X. Chemically recyclable polymers: a circular economy approach to sustainability. *Green Chem.* **19**, 3692–3706 (2017).

31. Parthasarathi, R., Romero, R. A., Redondo, A. & Gnanakaran, S. Theoretical study of the remarkably diverse linkages in lignin. *J. Phys. Chem. Lett.* **2**, 2660–2666 (2011).

32. Yu, D.-G., Li, B.-J. & Shi, Z.-J. Exploration of new C−O electrophiles in cross-coupling reactions. *Acc. Chem. Res.* **43**, 1486–1495 (2010).

33. Kärkäs, M. D., Matsuura, B. S., Monos, T. M., Magallanes, G. & Stephenson, C. R. J. Transition-metal catalyzed valorization of lignin: the key to a sustainable carbon-neutral future. *Org. Biomol. Chem.* **14**, 1853–1914 (2016).

34. Nichols, J. M., Bishop, L. M., Bergman, R. G. & Ellman, J. A. Catalytic C−O bond cleavage of 2-aryloxy-1-arylethanols and its application to the depolymerization of lignin-related polymers. *J. Am. Chem. Soc.* **132**, 12554–12555 (2010).

35. vom Stein, T. et al. Bond cleavage in lignin model substrates. *Angew. Chem. Int. Ed. Engl.* **54**, 5859–5863 (2015).

36. vom Stein, T., Weigand, T., Merkens, C., Klankermayer, J. & Leitner, W. Trimethylenemethane-ruthenium(II)-triphos complexes as highly active catalysts for catalytic C−O bond cleavage reactions of lignin model compounds. *ChemCatChem* **5**, 439–441 (2013).

37. Garrison, M. D. & Harvey, B. G. Bio-based hydrophobic epoxy-amine networks derived from renewable terpenoids. *J. Appl. Polym. Sci.* **133**, 43621 (2016).

38. De Vincenzi, M., Stammati, A., De Vincenzi, A. & Silano, M. Constituents of aromatic plants: carvacrol. *Fitoterapia* **75**, 801–804 (2004).

39. Costa, V. V., da Silva Rocha, K. A., Kozhevnikov, I. V., Kozhevnikova, E. F. & Gusevskaya, E. V. Heteropoly acid catalysts for the synthesis of

fragrance compounds from biorenewables: isomerization of limonene oxide. *Catal. Sci. Technol.* **3**, 244–250 (2013).

40. Mauck, J. R. et al. Preparation and characterization of highly bio-based epoxy amine thermosets derived from lignocellulosics. *Macromol. Chem. Phys.* **218**, 1700013 (2017).

41. Sutton, C. A. et al. Novel biobased furanic diols as potential alternatives to BPA: synthesis and endocrine activity screening. *ACS Sustain. Chem. Eng.* **8**, 18824–18829 (2020).

42. Huber, G. W., Iborra, S. & Corma, A. Synthesis of transportation fuels from biomass: chemistry, catalysts, and engineering. *Chem. Rev.* **106**, 4044–4098 (2006).

43. Coetzee, J. et al. Homogeneous catalytic hydrogenation of amides to amines. *Chem. Eur. J.* **19**, 11039–11050 (2013).

44. Çelenligil-Çetin, R., Watson, L. A., Guo, C., Foxman, B. M. & Ozerov, O. V. *Organometallics* **24**, 186–189 (2005).

45. Eze, W. U. et al. Plastics waste management: a review of pyrolysis technology. *Clean Technol. Recycl.* **1**, 50–69 (2021).

46. Kijo-Kleczkowska, A. & Gnatowski, A. Recycling of plastic waste, with particular emphasis on thermal methods – review. *Energies* **15**, 2114–2135 (2022).

47. Rani, M., Choudhary, P., Krishnan, V. & Zafar, S. A review on recycling and reuse methods for carbon fiber/glass fiber composites waste from wind turbine blades. *Compos. B Eng.* **215**, 108768–108783 (2021).

48. Gonçalves, R. M., Martinho, A. & Oliveira, J. P. Recycling of reinforced glass fibers waste: current status. *Materials* **15**, 1596–1614 (2022).

Acknowledgements

We thank the entire CETEC consortium—in particular M. E. Birkbak from Vestas Wind System A/S, S. C. A. Lauth, M. Schrötz and L. O. Meyer from the Olin Corporation, as well as E. Damgaard-Møller and A. S. Donslund from the Danish Technological Institute—for valuable discussions, epoxy resin/composite material samples and assistance regarding analytics. We thank M. Ceccato from Aarhus University for measurement of XPS spectra and S. S. Pedersen from Aarhus University for procuring a carbon fibre composite sample. We also thank CSCAA for the computing hours provided for the DFT study. We are deeply grateful for financial support by Innovation Fund Denmark (grant no. 0224-00072B), Carlsberg Foundation (grant no. CF18-1101), Danish National Research Foundation (grant nos. DNRF118 and DNRF-93), Novo Nordisk Foundation research infrastructure grant for AXIA (no. NNF19OC0055801) and Aarhus University. Support from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 862179 and Marie Sklodowska-Curie grant agreement no. 859910 is also gratefully acknowledged. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use that may be made of the information contained therein.

Author information

Authors and Affiliations

1. Department of Chemistry and Interdisciplinary Nanoscience Center, Aarhus University, Aarhus, Denmark

   Alexander Ahrens, Andreas Bonde, Hongwei Sun, Nina Kølln Wittig, Hans Christian D. Hammershøj, Gabriel Martins Ferreira Batista, Henrik Birkedal & Troels Skrydstrup

2. Danish Technological Institute, Aarhus, Denmark

   Andreas Sommerfeldt & Simon Frølich

Contributions

Conceptualization, writing and revising of the original draft were carried out by A.A. and T.S. Experimental design was carried out by A.A. Experimental investigation was carried out by A.A., A.B. and H.S. A.A. and T.S. supervised and directed research. Analytics and evaluation of recovered fibres were conducted by N.K.W., H.B., A.S. and S.F. X-ray crystallographic investigations were conducted by H.C.D.H. DFT studies were conducted by G.M.F.B. Preparation of clear-cast epoxy samples and mass analysis on remaining fraction were carried out by A.S. Funding acquisition was carried out by T.S. and S.F. All authors reviewed the final manuscript.

Corresponding authors

Correspondence to Alexander Ahrens or Troels Skrydstrup.

Ethics declarations

Competing interests

A.A., T.S., A.S. and S.F. are inventors on provisional patent application no. EP22156129, submitted by Aarhus University, which covers the transition-metal-catalysed disassembly of epoxy-based, fibre-reinforced, polymer composites. T.S. is co-owner of SyTracks A/S, which commercializes CO tubes.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

[**Supplementary Information**]

[**Peer Review File**]

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Ahrens, A., Bonde, A., Sun, H. *et al.* Catalytic disconnection of C–O bonds in epoxy resins and composites. *Nature* **617**, 730–737 (2023). https://doi.org/10.1038/s41586-023-05944-6

- Received: 23 August 2022
- Accepted: 13 March 2023
- Published: 26 April 2023

- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-05944-6

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

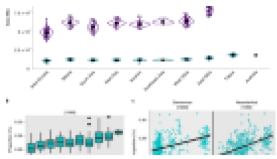
Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

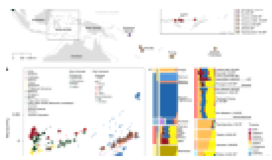
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-05944-6



Widespread Denisovan ancestry in Island Southeast Asia but no evidence of substantial super-archaic hominin admixture

22 March 2021

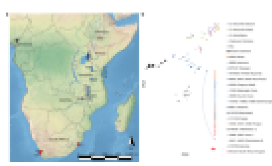
João C. Teixeira, Guy S. Jacobs, … Kristofer M. Helgen



Ancient genomes from the last three millennia support multiple human dispersals into Wallacea

09 June 2022

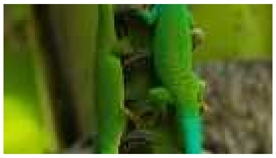
Sandra Oliveira, Kathrin Nägele, … Mark Stoneking



Ancient DNA and deep population structure in sub-Saharan African foragers

23 February 2022

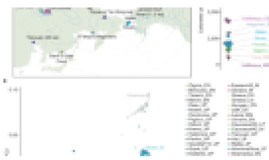
Mark Lipson, Elizabeth A. Sawchuk, … Mary E. Prendergast



[The Andaman day gecko paradox: an ancient endemic without pronounced phylogeographic structure]

16 July 2020

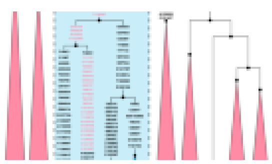
Ashwini V. Mohan, Pablo Orozco-terWengel, … Miguel Vences



[The population history of northeastern Siberia since the Pleistocene]

05 June 2019

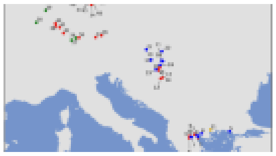
Martin Sikora, Vladimir V. Pitulko, … Eske Willerslev



[African mitochondrial haplogroup L7: a 100,000-year-old maternal human lineage discovered through reassessment and new sequencing]

24 June 2022

Paul A. Maier, Göran Runfeldt, … Miguel G. Vilar



Ancient mitochondrial diversity reveals population homogeneity in Neolithic Greece and identifies population dynamics along the Danubian expansion axis

05 August 2022

Nuno M. Silva, Susanne Kreutzer, … Christina Papageorgopoulou



Searching for the roots of the first free African American community

26 November 2020

Beatriz Martínez, Filipa Simão, … Leonor Gusmão



Peopling of the Americas as inferred from ancient genomics

16 June 2021

Eske Willerslev & David J. Meltzer

- Article
- Open Access
- Published: 17 May 2023

A weakly structured stem for human origins in Africa

- Aaron P. Ragsdale,
- Timothy D. Weaver,
- Elizabeth G. Atkinson,
- Eileen G. Hoal,
- Marlo Möller,
- Brenna M. Henn &
- …
- Simon Gravel

Show authors

Nature **volume 617**, pages 755–763 (2023)

- 65k Accesses
- 1140 Altmetric
- Metrics details

Abstract

Despite broad agreement that *Homo sapiens* originated in Africa, considerable uncertainty surrounds specific models of divergence and migration across the continent[1]. Progress is hampered by a shortage of fossil and genomic data, as well as variability in previous estimates of divergence times[1]. Here we seek to discriminate among such models by considering linkage disequilibrium and diversity-based statistics, optimized for rapid, complex demographic inference[2]. We infer detailed demographic models for populations across Africa, including eastern and western representatives, and newly sequenced whole genomes from 44 Nama (Khoe-San) individuals from southern Africa. We infer a reticulated African population history in which present-day population structure dates back to Marine Isotope Stage

5. The earliest population divergence among contemporary populations occurred 120,000 to 135,000 years ago and was preceded by links between two or more weakly differentiated ancestral *Homo* populations connected by gene flow over hundreds of thousands of years. Such weakly structured stem models explain patterns of polymorphism that had previously been attributed to contributions from archaic hominins in Africa[2,3,4,5,6,7]. In contrast to models with archaic introgression, we predict that fossil remains from coexisting ancestral populations should be genetically and morphologically similar, and that only an inferred 1–4% of genetic differentiation among contemporary human populations can be attributed to genetic drift between stem populations. We show that model misspecification explains the variation in previous estimates of divergence times, and argue that studying a range of models is key to making robust inferences about deep history.

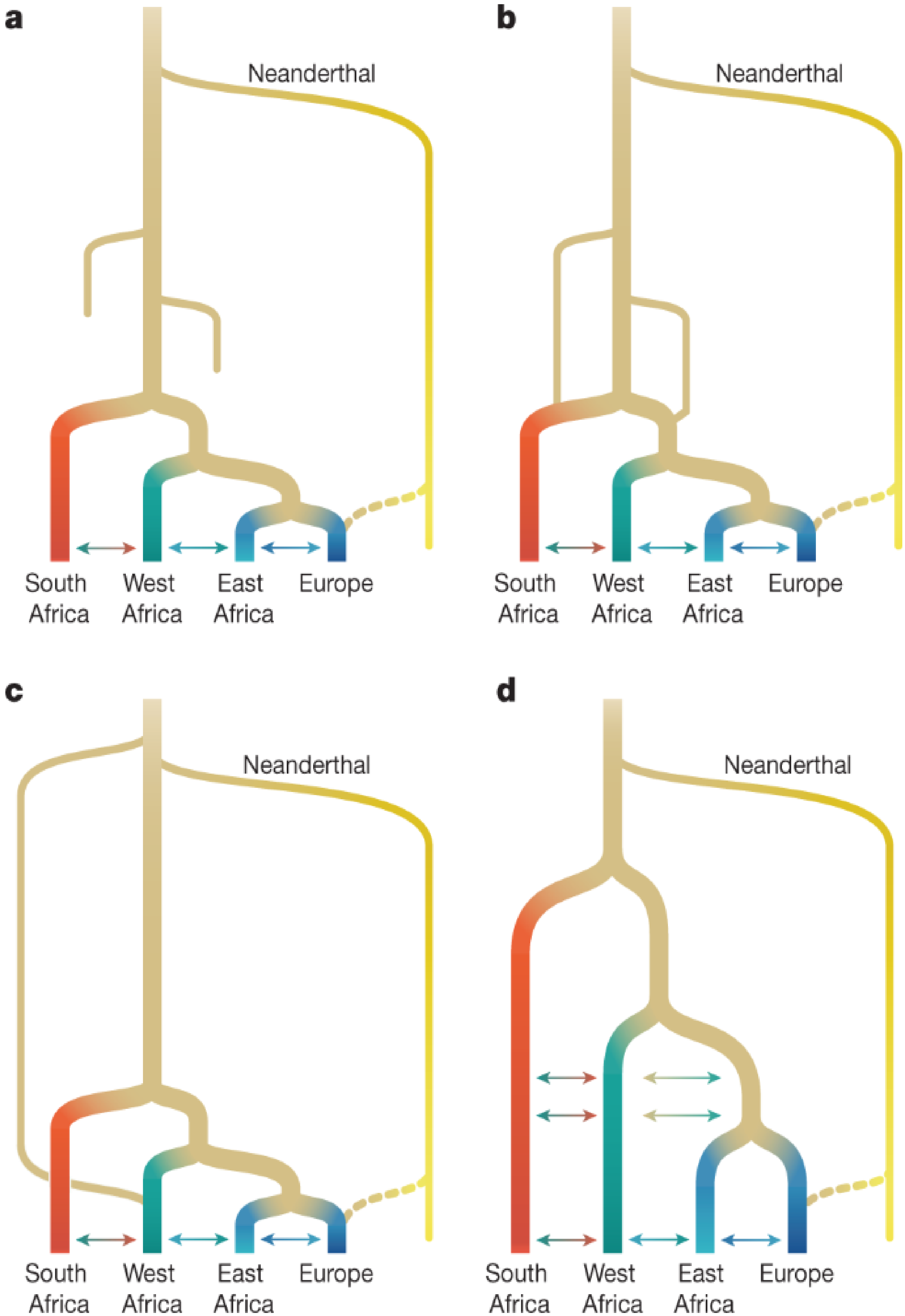
Main

Decades of study of human genome variation have suggested a predominantly tree-like model of recent population divergence from a single ancestral population in Africa. It has been difficult to reconcile this finding with the fossil and archaeological records of human occupation across the vast African continent. For example, fossils such as those from the sites of Jebel Irhoud in Morocco[8], Herto in Ethiopia[9] and Klasies River in South Africa[10] demonstrate that derived *Homo sapiens* anatomical features were found across the continent 300−100 thousand years ago (ka). Archaeological sites from the Middle Stone Age, of which some have been associated with *H. sapiens*, are also widely distributed across Africa. It is unclear whether these fossils and archaeological sites represent populations that contributed to contemporary *H. sapiens* as population precedents or were local 'dead ends'. Attempts to reconcile genetic and palaeoanthropological data include proposals[11,12,13] for a pan-African origin of *H. sapiens* in which populations in many regions of the continent contributed to the formation of *H. sapiens* beginning at least 300 ka.

Genetic models have been hampered in their contribution to this discussion because they primarily assume (or, at least, have been tested under) a tree-like model of isolation with migration. Alternative theoretical scenarios have been proposed, such as stepping-stone models[14] or population coalescence

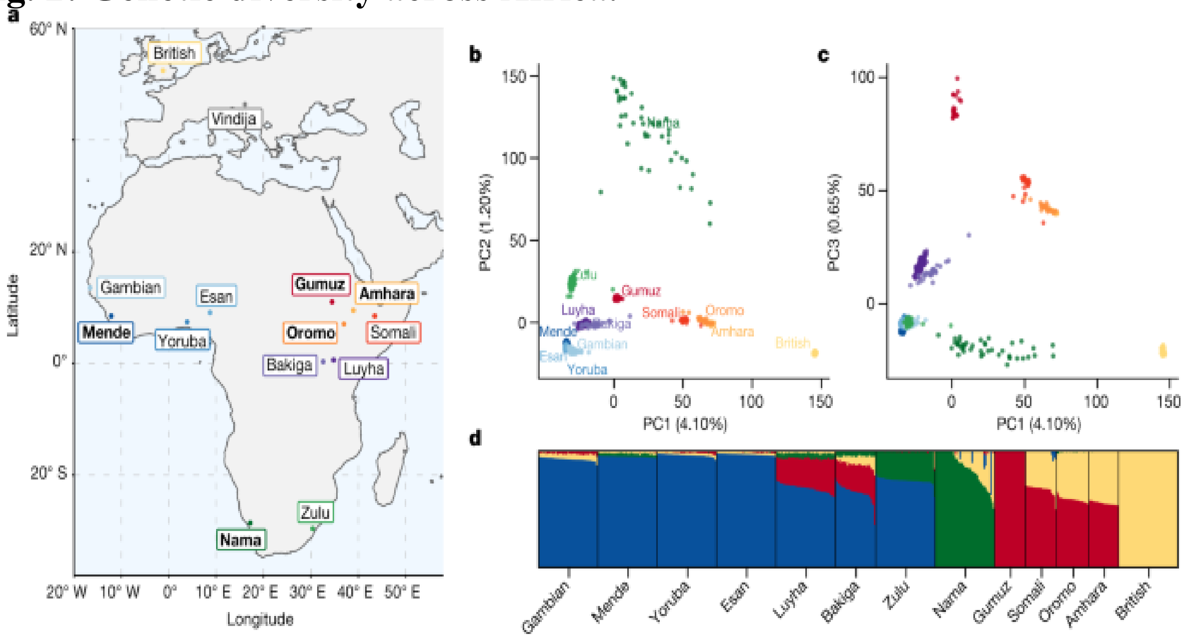
and fragmentation[13], but these approaches are more challenging to interpret and fit to data. However, new population-genetic tools now allow for inference on the basis of tens to hundreds of genomes from multiple populations and for greater model complexity[2,15,16]. Inspired by evidence for Neanderthal admixture with humans in Eurasia, several studies have shown that introducing an archaic hominin ‘ghost’ population contributing to African populations in the period surrounding the out-of-Africa migration event substantially improves the description of genetic data relative to single-origin models, mostly in western Africa[2,3,4,5,6,7], but also in southern[4,6] and central African[4,5,6,17] populations. This has driven speculation about the geographical range of the ghost population, possible links to specific fossils and the possibility of finding ancient DNA evidence[17]. However, these studies share two weaknesses. First, they contrast only a single-origin model with an archaic hominin admixture model, leaving out other plausible models[1] (Fig. 1). Second, they focus on a small subset of African diversity, either because of small sample sizes (2–5 genomes) or because they rely on data from the 1000 Genomes Project[18], which was limited to populations of recent West African or Bantu-speaking ancestry (Fig. 2). Ancient DNA from Eurasia has helped to clarify early human history outside Africa, but there is no comparably ancient DNA to elucidate early history in Africa[19].

Fig. 1: Proposed conceptual models of early human history in Africa.


a
Neanderthal
South Africa
West Africa
East Africa
Europe
b
Neanderthal
South Africa
West Africa
East Africa
Europe
c
Neanderthal
South Africa
West Africa
East Africa
Europe
d
Neanderthal
South Africa
West Africa
East Africa
Europe

a, Recent expansion. **b**, Recent expansion with regional persistence. **c**, Archaic admixture. **d**, African multiregional. The models have been designed to translate models from the palaeoanthropological literature into genetically testable demographic models (ref. [1] and Supplementary Information section 3). These parameters were then fitted to genetic data.

Fig. 2: Genetic diversity across Africa.



a, Selected populations from the 1000 Genomes Project and the African Diversity Reference Panel[18,20] illustrate diversity from western, eastern and southern Africa. We chose representative ethnic groups from each region (bold labels) to build parameterized models, including the newly genetically sequenced Nama populations from South Africa, Mende from Sierra Leone, Gumuz, Oromo and Amhara from Ethiopia, British individuals and a Neanderthal from Vindija Cave, Croatia. **b**,**c**, Principal component analysis highlights the range of genetic divergence anchored by western African, Nama, Gumuz and British individuals between principal components (PC) 1 and 2 (**b**), and 1 and 3 (**c**). Percentages show variance explained by each principal component. Colours represent the groups shown in bold in **a**. **d**, ADMIXTURE analysis using $K = 4$ principal components reveals signatures of recent gene flow in Africa that reflect colonial-period migration into the Nama, back-to-Africa gene flow among some Ethiopians, and Khoe-San admixture in the Zulu population.

We therefore aim to discriminate between a broader set of demographic models by studying the genomes of contemporary populations. We take as our starting point four models (single-population expansion, single-population expansion with regional persistence, archaic hominin admixture and multi-regional evolution; Fig. 1) using 290 genomes of individuals from southern, eastern and western Africa, as well as Eurasia. By including geographically and genetically diverse populations across Africa, we infer demographic models that explain more features of genetic diversity in more populations than previously reported. These analyses confirm the inadequacy of tree-like models and provide an opportunity to directly evaluate a wide range of alternative models.

We inferred detailed demographic histories using 4x–8x whole-genome sequencing data for four diverse African populations, comprising the Nama (Khoe-San from South Africa, newly presented here; see Supplementary Information section 1.2 for ethical and practical aspects of participant recruitment), the Mende (from Sierra Leone; from phase 3 of the 1000 Genomes Project[18]), the Gumuz (recent descendants of a hunter-gatherer group from Ethiopia[20,21]) and eastern African agriculturalists (Amhara and Oromo from Ethiopia[20]). The Amhara and Oromo populations, despite speaking distinct Afro-Asiatic languages, are highly genetically similar[21,22] so we combined the two groups for a larger sample size (Fig. 2). We also included British individuals from the 1000 Genomes Project in our demographic models as a representative source of back-to-Africa gene flow and recent colonial admixture in South Africa. Finally, we used a high-coverage ancient Neanderthal genome from Vindija Cave in Croatia[23] to account for gene flow from Neanderthals into people from outside Africa, and gauge the relative time depth of divergence, assuming that Neanderthals diverged 550 ka from a common stem. We computed one- and two-locus statistics for which the expectation within and across populations can be computed efficiently and that are well suited for both low- and high-coverage genomes[2,24]. Using a maximum-likelihood inference framework, we then fitted to these statistics a family of parameterized demographic models that involve population splits, size changes, continuous and variable migration rates and punctuated admixture events, to learn about the nature of the population structure over the past million years.

A Late Pleistocene common ancestry

We started with a model of geographical expansion from a single ancestral, unstructured source followed by migration between populations, without allowing for a contribution from an African archaic hominin lineage (Fig. 1a) or population structure before the expansion (Fig. 1d). As expected[2], this first model was a poor fit to the data qualitatively (Supplementary Fig. 10) and quantitatively (log-likelihood (LL) ≈ −189,300; Supplementary Table 3). We next explored a suite of parameterized models in which population structure predates the differentiation of contemporary groups (Supplementary Information section 3). Depending on the parameters, these encompassed models allowing for ancestral reticulation, such as fragmentation-and-coalescence or meta-population models (Fig. 1b), archaic hominin admixture (Fig. 1c) and African multi-regionalism (Fig. 1d). The recent expansion and the African multi-regional models (Fig. 1a,d) have the same topology, so interpretation of the model depends on the specified or inferred divergence times.

Regardless of the model choice for early epochs, maximum-likelihood inference of human demographic history for the past 150 kyr was remarkably robust. In a reticulated model, we use 'divergence' between populations to mean the time of their most recent shared ancestry. The earliest divergence among contemporary human populations differentiates the southern African Nama population from the other African groups at 110−135 ka, with low to moderate levels of subsequent gene flow (Table 1). In none of the high-likelihood models that we explored was the divergence between Nama and other populations earlier than around 140 ka. We conclude that geographical patterns of contemporary *H. sapiens* population structure probably arose during MIS 5. Although we do find evidence for earlier population structure in Africa, contemporary populations cannot be easily mapped onto the more ancient 'stem' groups because only a small proportion of drift between contemporary populations can be attributed to drift between stems (Fig. 4, Supplementary Information section 5.2 and Supplementary Figs. 16–19).

Table 1 Migration and divergence parameters from best-fit models

Given this consistency in inferred recent history and the numerical challenge of optimizing a large number of parameters, we fixed several parameters related to recent population history to focus on more-ancient events (Supplementary Information section 3.1). These parameters were ones supported by multiple genetic and archaeological studies[25]. Fixed parameters included the time of divergence between western and eastern African populations, set to 60 ka, just before the split of Eurasians and East Africans at 50 ka. We also fixed the amount of admixture from Neanderthals to the European population directly after the out-of-Africa migration to 1.5% at 45 ka.

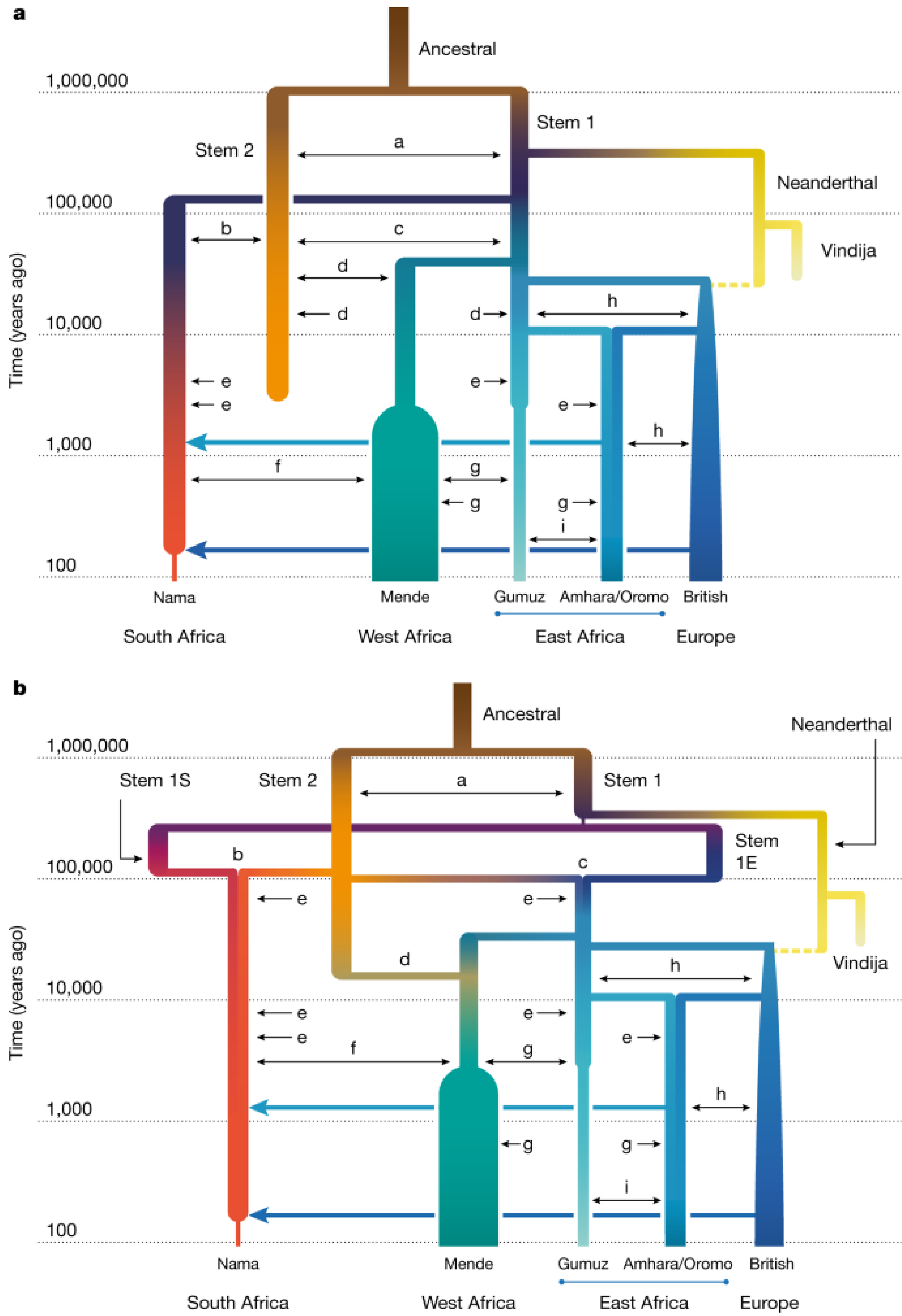
We quantify the migration rates of populations after their divergence at around 120 ka. Before the agropastoralist expansion 5 ka, migration between the ancestors of the Nama and other groups is an order of magnitude weaker than that observed between western and eastern Africans (Table 1). All models infer relatively high gene flow between eastern and western Africa ($m \approx 2 \times 10^{-4}$, the constant proportion of migrant lineages per generation since their divergence 60 ka). We further find that back-to-Africa gene flow at the beginning of the Holocene epoch primarily affected the ancestors of the Ethiopian agricultural populations[26], comprising almost 65% of their genetic ancestry. We observe considerable gene flow from the Amhara and Oromo into the Nama, a signal that is probably a proxy for the movement of eastern African caprid (goat) and cattle pastoralists[27,28], here estimated to constitute a 25% ancestry contribution 2 ka. Although this gene flow is not apparent from the ADMIXTURE plot (Fig. 2), the ancestry is probably grouped into the Khoe-San component, which has drifted appreciably from its ancestral eastern African source. Colonial-period admixture from Europeans into the Nama was estimated at 15%, similar to proportions suggested by ADMIXTURE (Fig. 2).

A weakly structured stem within Africa

To account for the population structure before 135 ka, three of our four models allowed for two or more stem populations, which could diverge either before or after the split from the Neanderthals. We considered models both with and without migration between these stem populations, and in both cases we tested two different types of gene flow during the expansion

phase, as illustrated in Supplementary Fig. 6: in the first, one of the stem population expands (splits into contemporary populations), followed by continuous symmetric migration with the other stem population(s); in the second, one or more of the stem populations expands, with instantaneous 'pulse' (merger) events from the other stem population, so that recent populations are formed by mergers of multiple ancestral populations. Depending on the parameter values, this scenario encompasses archaic hominin introgression and fragmentation-and-coalescence models (such as Fig. 1b,c). For many parameters, confidence intervals based on bootstrapping are relatively narrow (Supplementary Tables 3–7), reflecting an informative statistical approach. However, model assumptions have a greater effect on parameter estimates (and thus real uncertainty). To convey the uncertainty in the models, we highlight features of the two inferred models with high likelihoods. These are referred to as the multiple-merger and the continuous-migration models. Both allow for migration between stem branches, but differ primarily in the timing of the early divergence of stem populations and their relative effective population size (N_e) (Fig. 3). The two models also differ in the mode of divergence, with the multiple-merger model featuring a population reticulation (that is, loops in the population graph; Fig. 1b) during the Middle Pleistocene epoch (780 ka to 130 ka).

Fig. 3: A weakly structured stem best describes two-locus statistics.

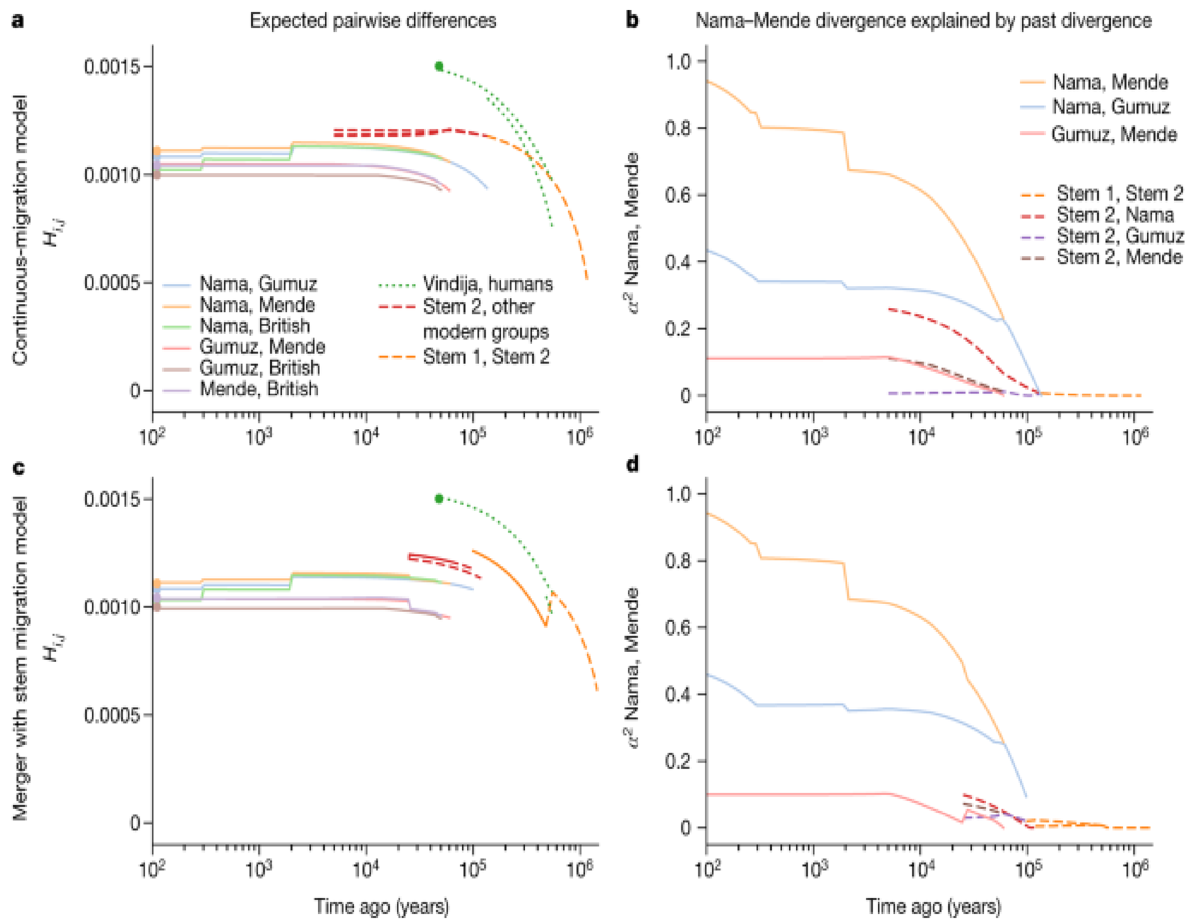


a
Ancestral
1,000,000
100,000
10,000
1,000
100
Time (years ago)
Stem 1
Stem 2
Neanderthal
Vindija
a
b
c
d
e
f
g
h
i
Nama
Mende
Gumuz
Amhara/Oromo
British
South Africa
West Africa
East Africa
Europe
b
Ancestral
Neanderthal
1,000,000
100,000
10,000
1,000
100
Time (years ago)
Stem 1S
Stem 2
Stem 1
Stem 1E
Vindija
Nama
Mende
Gumuz
Amhara/Oromo
British
South Africa
West Africa
East Africa
Europe

a,**b**, In the two best-fitting parameterizations of early population structure, continuous migration (**a**) and multiple mergers (**b**), models that include ongoing migration between stem populations outperform those in which stem populations are isolated. Most of the recent populations are also connected by continuous, reciprocal migration that is indicated by double-headed arrows (labels matched to migration rates and divergence times in Table 1). These migrations last for the duration of the coexistence of contemporaneous populations with constant migration rates over those intervals. The merger-with-stem-migration model (**b**, with LL = −101,600) outperformed the continuous-migration model (**a**, with LL = −115,300). Colours are used to distinguish overlapping branches. The letters a–i represent continuous migration between pairs of populations, as described in Table 1.

Allowing for continuous migration between the stem populations substantially improves the fits relative to zero migration between stems (LL ≈ −101,600 compared with −107,700 in the merger model, Supplementary Tables 6 and 7; and LL ≈ −115,300 versus −126,500 in the continuous migration model, Supplementary Tables 4 and 5). With continuous migration between stems, population structure extends back to more than 1 million years ago (Table 1). Migration between the stems in these models is moderate, with a fraction of migrant lineages (m) in each generation estimated as $m = 6.3 \times 10^{-5}$–1.3×10^{-4}. For comparison, this is similar to the inferred migration rates between connected contemporary populations over the past 50 ka (Table 1). This ongoing (or at least, periodic) gene flow qualitatively distinguishes these models from previously proposed archaic hominin admixture models (Fig. 1c), as the early branches remain closely related and each branch contributes large amounts to all contemporary populations (Fig. 4). Because of this relatedness, only 1–4% of genetic differentiation among contemporary populations can be traced back to this early population structure (Supplementary Information section 5.2).

Fig. 4: Structure among stems is weak and present-day structure is generally recent.



a–**d**, From the best-fit models of our two parameterizations (**a**,**b**, continuous migration; **c**,**d**, merger with stem migration), we predicted differentiation and shared drift between populations at past time points. **a**,**c**, We computed expected pairwise differences $H_{i,j}$ between individuals sampled from populations i and j existing at time t. **b**,**d**, To understand how drift between stems explains contemporary structure, we computed the proportion α^2 of drift between pairs of sampled contemporary populations (here the Nama and Mende) that aligns with drift between past populations (see Supplementary Information section 5.2 for details and additional comparisons in Supplementary Figs. 16–19). Both models infer deep population structure with modest contributions to contemporary genetic differentiation. Most present-day differentiation dates back to the past 100 kyr.

Under the continuous-migration model, one of the two stems (stem 1) diverges into lineages leading to contemporary populations in western, southern and eastern Africa, and the other (stem 2) contributes variable

ancestry to those populations. This migration from stem 2 is highest with the Mende ($m = 1.6 \times 10^{-4}$) compared with the Nama and populations from eastern Africa ($m = 5.9 \times 10^{-5}$ and 3.1×10^{-5}, respectively), with migration allowed to occur until 5 ka. A sampled lineage from the Nama, Mende and Gumuz have probabilities of being in stem 2 at the time of stem 1 expansion (135 ka) of approximately 0.145, 0.20 and 0.130, respectively, although these probabilities change over time, precluding the notion of a fixed admixture proportion.

By contrast, in the multiple-merger model, stem populations merge with varying proportions to form the different contemporary groups. We observe a sharp bottleneck in stem 1 down to $N_e = 100$ after the split of the Neanderthal branch. This represents the lower bound allowed in our optimization (an N_e of 100), although the size of this bottleneck is poorly constrained (95% confidence interval 100–851). After a long period of exchange with stem 2, stem 1 then fractures into stem 1E and stem 1S at 478 ka. The timing of this divergence was also poorly constrained (95% confidence interval 276– 478 ka). These populations evolve independently until 119 ka (101−125 ka) when stem 1S and stem 2 combine to form the ancestors of the Nama, with proportions of 30% and 70%, respectively. Similarly, stem 1E and stem 2 combine in equal proportions (50% each) to form the ancestors of the western and eastern Africans (and thus also all individuals who later disperse during the out-of-Africa event). Finally, the Mende receive a large additional pulse of gene flow from stem 2, replacing 19% (18–21%) of their population 25 ka (22−26 ka). The later stem 2 contribution to the western African Mende resulted in better model fits ($\Delta LL \approx 60{,}000$). This may indicate that an ancestral stem 2 population occupied western or central Africa, broadly speaking. The differing proportions in the Nama and eastern Africans may also indicate a geographical separation of stem 1S in southern Africa and stem 1E in eastern Africa.

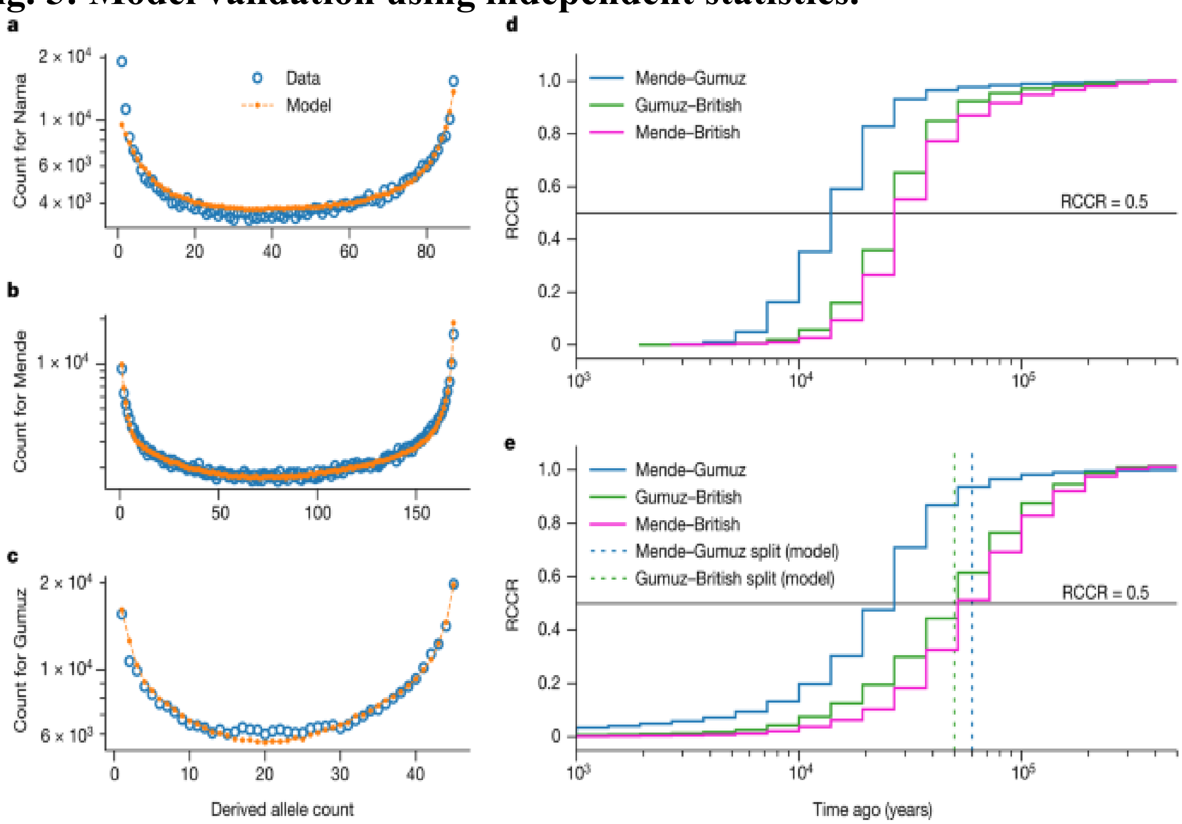
To assess the robustness of the inferred models to analysis and reference population choices, Supplementary Information sections 6 and 7 include reanalyses with changes in the European and West African populations, as well as the recombination maps, filtering strategies and parameter optimization strategies. Although we find some differences in the inferred

parameters (see Supplementary Information sections 7.1.1 and 7.2), the best-fit models across all reanalyses are quantitatively consistent.

Reconciling lines of genetic evidence

Previous studies have found support for archaic hominin admixture in Africa using two-locus statistics[2,17], conditional site frequency spectra (cSFS)[7] and the reconstruction of gene genealogies[16]. However, none of these studies considered a weakly structured stem. We validated our inferred models with additional independent approaches. We find that the observed cSFS (conditional on the derived allele being carried in the Neanderthal sample) is well described by the merger model (Fig. 5a–c and Supplementary Figs. 20–23), even though this statistic was not used in the fit. Our best-fit models outperform archaic hominin admixture models fitted directly to the cSFS (for example, compare with figure 1 in ref. [7]). Specifically, it is the addition of migration between stems that results in a qualitative improvement of the agreement (compare Supplementary Figs. 22 and 23).

Fig. 5: Model validation using independent statistics.



a–**c**, Using our best-fit models, we simulated expected cSFS and compared the simulated spectra to those observed from the data. Our inferred models provide a good fit to the data, even though this summary was not used in our inference. Across the three populations (**a**, Nama; **b**, Mende; **c**, Gumuz), ancestral-state misidentification was consistently inferred to be 1.5−1.7% for intergenic loci (Supplementary Information section 6.2.2). **d**,**e**, We used Relate[16] to reconstruct genome-wide genealogies, which we used to estimate coalescence-rate trajectories and cross-coalescence rates between pairs of populations. Although coalescence-rate distributions are informative about past evolutionary processes, interpretation can be hindered by migration and population structure, and translating RCCR curves into population divergence times is especially prone to misinterpretation. **d**, Real data; **e**, our model. In our model, the Mende–Gumuz split occurs before the Gumuz–British split. However, the model also predicts a recent elevated Mende–Gumuz RCCR. This pattern, also observed in the data, does not indicate that the Mende and Gumuz split more recently than the Gumuz and British populations.

We used the software Relate[16] to infer the distribution in the coalescence rates over time in both real data and data simulated from our inferred models. Many previous studies have found a reduction of coalescence rates between 1 million years ago and 100 ka in humans and thus inferred an increase in N_e during the same period[29]. This increase in inferred N_e could be attributable to either an increase in population size or to ancestral population structure during the Middle Pleistocene[30]. All the models, including the single-origin model, recapitulate an inferred ancestral increase in N_e between 100 ka and 1 million years ago (Supplementary Fig. 26 and Supplementary Information section 7.3.2). The single-origin model achieves this by an increase in N_e during that period, whereas the best-fit models recapitulate this pattern without corresponding changes in population size.

Relative cross-coalescence rates (RCCRs) have recently been used to estimate divergence between pairs of populations, as measured by the rate of coalescence between two groups divided by the mean within population coalescence. Simulations of RCCR accuracy, however, focus on a clean split between populations, whereby groups diverge without subsequent gene flow. Published estimates[25] of the earliest human divergences with RCCR, which

range from 150 ka to 100 ka, may be substantially biased when compared with more-complex models with gene flow as inferred here. We find that midpoint estimates of RCCR are poor estimates for population divergence, often underestimating divergence time by 50% or more (for example, Mende versus Gumuz is about 15 ka compared with a true divergence of 60 ka), and recent migration can lead to the misordering of divergence events (Fig. 5e). We suggest that RCCR analyses that do not fit multiple parameters, including gene flow, should be interpreted with caution.

Other studies[1,25] have fitted tree-like demographic models to African populations using distributions of allele frequencies or related statistics, finding inconsistent divergence times, some of which are older than those we find here. In Supplementary Information section 7.4, we show that this discrepancy can be explained by model misspecification: if divergence is estimated by using an isolation with migration model with constant population sizes, but the correct model has ancient population growth or population structure, the divergence time in the inferred model is much earlier than in the correct model. Intuitively, growth or structure in the ancestral population will each increase coalescence times relative to a randomly mating population of constant size, so a model that assumes constant population sizes would require an older divergence time to fit the observed distribution of coalescence times and related statistics[31,32].

Discussion

Any attempt to build detailed models of human history is subject to model misspecification. This is true of previous studies, which often assumed that data inconsistent with a single-origin model should be explained by archaic hominin admixture. It is also true of this study. Although it is difficult to fully explore the space of plausible models of early human population structure, we sought to capture uncertainty in the model by exploring multiple parameterizations of early history. The best-fit models presented here include reticulation and migration between early human populations, rather than archaic hominin admixture from long-isolated branches (Fig. 1c). Elements of both recent expansion and African multiregionalism (Fig. 1a,d) feature in our best-fit models, as indicated in the recent time of

contemporary population divergence and the gene flow between disparate stems, respectively.

We cannot rule out the possibility that more-complex models involving additional stems, more-complex population structure, or hybrid models including both weak structure and archaic hominin admixture, may better explain the data. Because parameters related to the split time, migration rates and relative sizes of the early stems were variable across models, reflecting a degree of confounding among these parameters, we refrained from introducing additional branches associated with more parameters during that period. Rather than interpreting the two stems as representing well-defined and stable populations over hundreds of thousands of years, we interpret the weakly structured stem as consistent with a population fragmentation-and-coalescence model[13]. Other African populations, such as those from Central Africa, other Khoe-San groups or pre-Holocene ancient DNA samples, could further test our proposed models.

Formation of population structure in Africa

Our inferred models paint a consistent picture of the Middle to Late Pleistocene as a critical period of change, assuming that estimates from the recombination clock accurately relate to geological chronologies (Supplementary Information section 8). During the late Middle Pleistocene, the multiple-merger model indicates three major stem lineages in Africa, tentatively assigned to southern (stem 1S), eastern (stem 1E) and western/central Africa (stem 2). Geographical association was informed by the present population location with the greatest ancestry contribution from each stem. For example, stem 1S contributes 70% to the ancestral formation of the Khoe-San. The extent of the isolation 400 ka between stem 1S, stem 1E and stem 2 suggests that these stems were not proximate to each other. Although the length of isolation among the stems is variable across fits, models with a period of divergence, isolation and then a merger event (that is, a reticulation) out-performed models with bifurcating divergence and continuous gene flow.

A population reticulation involves multiple stems that contribute genetically to the formation of a group. One way in which this can happen is through the

geographical expansion of one or both stems. For example, if, during MIS 5, either stem 1S (Fig. 3b) from southern Africa moved northwards and thus encountered stem 2, or stem 2 moved from central–western Africa southwards into stem 1S, then we could observe disproportionate ancestry contributions from different stems in contemporary groups. We observed two merger events. The first, between stem 1S and stem 2, resulted in the formation of an ancestral Khoe-San population around 120 ka. The second event, between stem 1E and stem 2 about 100 ka, resulted in the formation of the ancestors of eastern and western Africans, including the ancestors of people outside Africa. Reticulated models do not have a unique and well-defined basal human population divergence. We suggest conceptualizing the events at 120 ka as the time of most recent shared ancestry among sampled populations. However, interpreting population divergence times in population genetics is always difficult, owing to the co-estimation of divergence time and subsequent migration; methods assuming clean and reticulated splits can infer different split dates (Supplementary Figs. 28 and 36). Therefore, in the literature, wide variation exists in estimates of divergence time[1,25].

Shifts in wet and dry conditions across the African continent between 140 ka and 100 ka may have promoted these merger events between divergent stems. Precipitation does not neatly track interglacial cycles in Africa, and heterogeneity across regions may mean that the beginning of an arid period in eastern Africa is conversely the start of a wet period in southern Africa[33]. The rapid rise in sea levels during the MIS 5e interglacial might have triggered migration inland away from the coasts, as has been suggested, for example, for the palaeo-Agulhas plain[34]. After these merger events, the stems subsequently fractured into subpopulations which persisted over the past 120 ka. These subpopulations can be linked to contemporary groups despite subsequent gene flow across the continent. For example, a genetic lineage sampled in the Gumuz has a probability of 0.7 of being inherited from the ancestral eastern subpopulation 55 ka, compared with a probability of 0.06 of being inherited from the southern subpopulation (see Table S8 for additional comparisons).

We also find that stem 2 continued to contribute to western Africans during the Last Glacial Maximum (26 ka to 20 ka), indicating that this gene flow

probably occurred in western and/or central Africa (Table 1). Such an interpretation is reinforced by differential migration rates between regions; that is, the gene flow from stem 2 to western Africans is estimated to be five times that of the rate to eastern Africans during this period. We performed a variety of validation tests to explore the sensitivity of our assumptions, including relaxing fixed parameters (Supplementary Information section 6). Most of the validation tests resulted in parameters similar to the models discussed above. However, one exception was the inferred out-of-Africa and eastern–western African divergences, which were 10–15 ka younger than our fixed parameters. These younger dates are at odds with the accepted timing of the out-of-Africa expansion that contributed to later human populations at approximately 50 ka, based on archaeological, climatic and fossil information[35,36,37,38]. Because the inference approach is unbiased in simulations, we interpret the free estimate for eastern African versus European divergence as reflecting our inclusion of only a single out-of-Africa population in the model, the lack of a nearby source for back-to-Africa gene flow, and other regionally complex parameters, rather than a systematic bias that may affect all parameters in the model. Older pan-African features of our inferred models are minimally affected by the choice of these fixed parameters (Supplementary Information section 7.2).

Contrasting ancestral structure models

Evidence for archaic hominin admixture in Eurasia has bolstered the plausibility of archaic hominin admixture having also occurred in Africa. Previous work that sought to explain patterns of polymorphism inconsistent with a single-origin model therefore focused on archaic hominin admixture as an alternative model, by referring to additional (ghost) branches required to fit the data as archaic[2,3,4,5,6,7] and assuming (or inferring) deep divergences. These perspectives have oriented interpretations of both genomic (for example, selection[39]) and fossil (such as the evolution of early *H. sapiens*[40]) data. Here we have shown that a weakly structured stem model better captures the apparently inconsistent patterns of polymorphisms.

Preferring models of a weakly structured stem to archaic-admixture models has a range of implications. First, with a weakly structured stem, there is no need to posit that an archaic hominin population in Africa remained

reproductively isolated from the ancestral human lineage for hundreds of thousands of years before the initiation of gene flow. Instead, there would simply have been continuous or recurrent contact between two or more groups present in Africa.

Second, there is evidence for both deleterious and adaptive archaic-hominin-derived alleles in contemporary genomes in the form of a depletion of Neanderthal ancestry in regulatory regions[41], or an increased frequency of archaic-hominin-related haplotypes such as at *EPAS1* among Tibetan people[42]. Under previous African archaic-hominin admixture models, the estimated 8–10% introgression rate is much higher than Neanderthal gene flow and would have plausibly been fertile ground for considerable selection for or against archaic-hominin-derived haplotypes[39]. By contrast, adaptation under a weakly structured stem would have occurred continuously over much longer periods. Polymorphism patterns that are inconsistent with the single-stem model predictions have been used to infer putative archaic admixed segments[3,7,17,39], negative selection against such segments[39] and pervasive positive selection[43]. However, such approaches are subject to large numbers of false positives in the presence of population structure with migration[41], and their interpretation should be re-examined in the light of a weakly structured stem model within Africa.

Third, multiple studies have shown a correspondence between phenotypic differentiation, usually assessed by measurements of the cranium, and genetic differentiation among human populations and between humans and Neanderthals[44,45,46] (see also Supplementary Information section 5.4). This correspondence potentially allows predictions of our model to be related to the fossil record. Some *H. sapiens* fossils, such as those from Iho Eleru in Nigeria (13 ka)[47], Ishango in the Democratic Republic of Congo (20–25 ka)[48] and Nazlet Khater in Egypt (35–40 ka)[49], have morphological features that may reflect recent gene flow from archaic hominins[47,48], and have been used in support of previously inferred archaic admixture scenarios[7,12,25]. The weakly structured stem model is not incompatible with archaic admixture having occurred in the ancestry of these fossils, but would imply, by contrast, that such individuals are unlikely to have contributed much ancestry to contemporary humans. The fossil record of Africa is sparse during the earlier time period of the stems ($\gtrsim$200 ka). Of the fossils that date

to this period, some are fairly similar overall in morphology to contemporary humans (for example, Omo 1 from Omo Kibish in Ethiopia[50,51]), whereas others are similar in some morphological features to contemporary humans (for example, Irhoud 1 from Jebel Irhoud in Morocco[8,52]); others are different enough in morphology to have been assigned to species other than *H. sapiens* (for example, DH1 from Dinaledi in South Africa[53,54]). If, as our model predicts, the genetic differences between the stems were similar to those among contemporary human populations, the most morphologically divergent fossils are unlikely to represent branches that contributed appreciably to contemporary human ancestries.

Methods

Data and sequencing

We generated a sequencing dataset by combining existing and newly recruited populations who are now part of the African Diversity Reference Panel (ADRP)[20,22], as well as the 1000 Genomes Project (1KGP) populations[18]. These included the Amhara, Bakiga, Gumuz, Nama (newly generated), Oromo, Somali and Zulu populations from the ADRP and ESN, GWD, LWK, MSL, YRI, CEU, GBR, CHB and PJL from 1KGP (these groups are defined elsewhere[18]). After filtering for relatedness and retaining Nama individuals with more than 70% estimated Khoe-San ancestry, we focused on data from 289 individuals, including 44 Nama. These were merged with the high-coverage Neanderthal genome from Vindija Cave[23]. We kept variants from regions that fell within the 1KGP strict-callability mask, overlapped with at least 100 continuously called base pairs in the Neanderthal genome and were annotated as intergenic. ADMIXTURE and principal component analyses were done on a subset of variants filtered to remove variants in high linkage disequilibrium (r^2 threshold of 0.1). Additional details on the data and sequencing are available in Supplementary Information section 1.

Linkage disequilibrium and diversity statistics

We used multi-population linkage disequilibrium and pairwise diversity statistics to fit parameterized demographic models to the data[2]. Unbiased linkage disequilibrium statistics were computed from all variants in retained intergenic regions[24], for pairs of variants separated by recombination distances $r = 5 \times 10^{-6}$–5×10^{-3} Morgans. These were assigned to 16 recombination distance bins, and average statistics were computed within each bin (Supplementary Information section 2.2). Expected statistics under each model were computed in Moments, which also performed likelihood-based parameter optimization. The cSFS were computed conditioned on the Vindija Neanderthal carrying the derived allele relative to the ancestral allele determined by a six-primate alignment[18].

Model specification and fitting

Model parameters include population sizes and size changes, split times, continuous migration rates and admixture times and proportions. The simplest model we tested was a bifurcating tree-like structure, allowing for subsequent migrations and recent known admixture events. To include ancestral population structure, we tested models that included multiple stem groups, each of which were allowed their own population size and could be connected by continuous migration. We tested multiple scenarios of early population structure, including long-lasting continuous migration between stem populations and scenarios of periods of isolation with subsequent merger events (Supplementary Information section 3). To avoid overfitting, we incrementally added complexity to our model optimization and we fixed a number of parameters that are constrained by historical records or are consistently estimated across multiple models and previous studies (Supplementary Information section 3.1). Likelihoods were computed using a composite multivariate Gaussian likelihood approach, and confidence intervals were estimated by refitting each model to 200 block-bootstrap replicate datasets (Supplementary Information section 3.2). We iteratively used gradient descent and L-BFGS-B optimization routines to fit each parameterized model (Supplementary Information section 3.3).

Gene genealogy reconstruction

We used Relate[16] to reconstruct genome-wide gene genealogies from the focal populations in the merged ADRP and 1KGP datasets (Supplementary Information section 4). From reconstructed genealogies, we computed coalescence rates within and between populations, which provide an estimate for effective population sizes over time and the relative cross-coalescence rates between pairs of populations. To compare reconstructed genealogies from data to model predictions, we used msprime[55,56] to simulate genomic data for equal numbers of samples for each population in our inferred models. We then applied Relate to these simulated datasets using the same mutation and recombination rates and generation time.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Nama sequencing data are available from the European Genome-Phenome Archive (EGA), accession number EGAD00001006198. Data access is permitted for non-commercial, population origins or ancestry research upon application to the South African Data Access Committee with appropriate institutional review board approval. The African Diversity Reference Panel can be found at accession EGAS00001000960.

Code availability

Code for the software used in this paper is found at the following locations: moments-LD (https://bitbucket.org/simongravel/moments), Demes (https://github.com/popsim-consortium/demes-python), Relate (https://myersgroup.github.io/relate/), msprime (https://github.com/tskit-dev/msprime) and tskit (https://github.com/tskit-dev/tskit).

References

1. Henn, B. M., Steele, T. E. & Weaver, T. D. Clarifying distinct models of modern human origins in Africa. *Curr. Opin. Genet. Dev.* **53**, 148–156 (2018).

2. Ragsdale, A. P. & Gravel, S. Models of archaic admixture and recent history from two-locus statistics. *PLoS Genet.* **15**, e1008204 (2019).

3. Plagnol, V. & Wall, J. D. Possible ancestral structure in human populations. *PLoS Genet.* **2**, e105 (2006).

4. Hammer, M. F., Woerner, A. E., Mendez, F. L., Watkins, J. C. & Wall, J. D. Genetic evidence for archaic admixture in Africa. *Proc. Natl Acad. Sci. USA* **108**, 15123–15128 (2011).

5. Hey, J. et al. Phylogeny estimation by integration over isolation with migration models. *Mol. Biol. Evol.* **35**, 2805–2818 (2018).

6. Lorente-Galdos, B. et al. Whole-genome sequence analysis of a pan African set of samples reveals archaic gene flow from an extinct basal population of modern humans into sub-Saharan populations. *Genome Biol.* **20**, 77 (2019).

7. Durvasula, A. & Sankararaman, S. Recovering signals of ghost archaic introgression in African populations. *Sci. Adv.* **6**, eaax5097 (2020).

8. Hublin, J.-J. et al. New fossils from Jebel Irhoud, Morocco and the pan-African origin of *Homo sapiens*. *Nature* **546**, 289–292 (2017).

9. White, T. D. et al. Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* **423**, 742–747 (2003).

10. Deacon, H. J. Two Late Pleistocene-Holocene archaeological depositories from the Southern Cape, South Africa. *S. Afr. Archaeol. Bull.* **50**, 121–131 (1995).

11. Stringer, C. The origin and evolution of *Homo sapiens*. *Phil. Trans. R. Soc. B* **371**, 20150237 (2016).

12. Scerri, E. M. L. et al. Did our species evolve in subdivided populations across Africa, and why does it matter? *Trends Ecol. Evol.* **33**, 582–594 (2018).

13. Scerri, E. M. L., Chikhi, L. & Thomas, M. G. Beyond multiregional and simple out-of-Africa models of human evolution. *Nat. Ecol. Evol.* **3**, 1370–1372 (2019).

14. Arredondo, A. et al. Inferring number of populations and changes in connectivity under the n-island model. *Heredity* **126**, 896–912 (2021).

15. Kamm, J., Terhorst, J., Durbin, R. & Song, Y. S. Efficiently inferring the demographic history of many populations with allele count data. *J. Am. Stat. Assoc.* **115**, 1472–1487 (2020).

16. Speidel, L., Forest, M., Shi, S. & Myers, S. R. A method for genome-wide genealogy estimation for thousands of samples. *Nat. Genet.* **51**, 1321–1329 (2019).

17. Hsieh, P. et al. Model-based analyses of whole-genome data reveal a complex evolutionary history involving archaic introgression in Central African Pygmies. *Genome Res.* **26**, 291–300 (2016).

18. The 1000 Genomes Project Consortium. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

19. Lipson, M. et al. Ancient DNA and deep population structure in sub-Saharan African foragers. *Nature* **603**, 290–296 (2022).

20. Gurdasani, D. et al. The African Genome Variation Project shapes medical genetics in Africa. *Nature* **517**, 327–332 (2015).

21. Gopalan, S. et al. Hunter-gatherer genomes reveal diverse demographic trajectories during the rise of farming in Eastern Africa. *Curr. Biol.* **32**, 1852–1860 (2022).

22. Pagani, L. et al. Tracing the route of modern humans out of Africa by using 225 human genome sequences from Ethiopians and Egyptians. *Am. J. Hum. Genet.* **96**, 986–991 (2015).

23. Prüfer, K. et al. A high-coverage Neandertal genome from Vindija Cave in Croatia. *Science* **358**, 655–658 (2017).

24. Ragsdale, A. P. & Gravel, S. Unbiased estimation of linkage disequilibrium from unphased data. *Mol. Biol. Evol.* **37**, 923–932 (2020).

25. Bergström, A., Stringer, C., Hajdinjak, M., Scerri, E. M. L. & Skoglund, P. Origins of modern human ancestry. *Nature* **590**, 229–237 (2021).

26. Molinaro, L. et al. West Asian sources of the Eurasian component in Ethiopians: a reassessment. *Sci. Rep.* **9**, 18811 (2019).

27. Henn, B. M. et al. Y-chromosomal evidence of a pastoralist migration through Tanzania to southern Africa. *Proc. Natl Acad. Sci. USA* **105**, 10693–10698 (2008).

28. Breton, G. et al. Lactase persistence alleles reveal partial East African ancestry of southern African Khoe pastoralists. *Curr. Biol.* **24**, 852–858 (2014).

29. Li, H. & Durbin, R. Inference of human population history from individual whole-genome sequences. *Nature* **475**, 493–496 (2011).

30. Mazet, O., Rodríguez, W., Grusea, S., Boitard, S. & Chikhi, L. On the importance of being structured: instantaneous coalescence rates and human evolution—lessons for ancestral population size inference? *Heredity* **116**, 362–371 (2016).

31. Momigliano, P., Florin, A.-B. & Merilä, J. Biases in demographic modeling affect our understanding of recent divergence. *Mol. Biol. Evol.* **38**, 2967–2985 (2021).

32. Shchur, V., Brandt, D. Y. C., Ilina, A. & Nielsen, R. Estimating population split times and migration rates from historical effective population sizes. Preprint at https://www.biorxiv.org/content/10.1101/2022.06.17.496540v1 (2022).

33. Blome, M. W., Cohen, A. S., Tryon, C. A., Brooks, A. S. & Russell, J. The environmental context for the origins of modern human diversity: a synthesis of regional variability in African climate 150,000–30,000 years ago. *J. Hum. Evol.* **62**, 563–592 (2012).

34. Marean, C. W. et al. in *Fynbos* (ed. Allsopp, N.) Ch. 8 (Oxford Univ. Press, 2014).

35. Fu, Q. et al. Genome sequence of a 45,000-year-old modern human from western Siberia. *Nature* **514**, 445–449 (2014).

36. Groucutt, H. S. et al. Rethinking the dispersal of *Homo sapiens* out of Africa. *Evol. Anthropol.* **24**, 149–164 (2015).

37. Prüfer, K. et al. A genome sequence from a modern human skull over 45,000 years old from Zlatý kůň in Czechia. *Nat. Ecol. Evol.* **5**, 820–825 (2021).

38. Beyer, R. M., Krapp, M., Eriksson, A. & Manica, A. Climatic windows for human migration out of Africa in the past 300,000 years. *Nat. Commun.* **12**, 4889 (2021).

39. Wall, J. D., Ratan, A., Stawiski, E. & GenomeAsia 100K Consortium. Identification of African-specific admixture between modern and archaic humans. *Am. J. Hum. Genet.* **105**, 1254–1261 (2019).

40. Grün, R. et al. Dating the skull from Broken Hill, Zambia, and its position in human evolution. *Nature* **580**, 372–375 (2020).

41. Petr, M., Pääbo, S., Kelso, J. & Vernot, B. Limits of long-term selection against Neandertal introgression. *Proc. Natl Acad. Sci. USA* **116**, 1639–1644 (2019).

42. Zhang, X. et al. The history and evolution of the Denisovan-EPAS1 haplotype in Tibetans. *Proc. Natl Acad. Sci. USA* **118**, e2020803118 (2021).

43. Schrider, D. R. & Kern, A. D. Soft sweeps are the dominant mode of adaptation in the human genome. *Mol. Biol. Evol.* **34**, 1863–1877

(2017).

44. Relethford, J. H. Craniometric variation among modern human populations. *Am. J. Biol. Anthropol.* **95**, 53–62 (1994).

45. Weaver, T. D., Roseman, C. C. & Stringer, C. B. Close correspondence between quantitative- and molecular-genetic divergence times for Neandertals and modern humans. *Proc. Natl Acad. Sci. USA* **105**, 4645–4649 (2008).

46. von Cramon-Taubadel, N. Congruence of individual cranial bone morphology and neutral molecular affinity patterns in modern humans. *Am. J. Phys. Anthropol.* **140**, 205–215 (2009).

47. Harvati, K. et al. The Later Stone Age calvaria from Iwo Eleru, Nigeria: morphology and chronology. *PLoS ONE* **6**, e24024 (2011).

48. Crevecoeur, I., Brooks, A., Ribot, I., Cornelissen, E. & Semal, P. Late Stone Age human remains from Ishango (Democratic Republic of Congo): new insights on Late Pleistocene modern human diversity in Africa. *J. Hum. Evol.* **96**, 35–57 (2016).

49. Crevecoeur, I. in *Modern Origins: A North African Perspective* (eds. Hublin, J. J. & McPherron, S. P.) 205–219 (Springer, 2012).

50. Day, M. H. Early *Homo sapiens* remains from the Omo River region of South-west Ethiopia: Omo human skeletal remains. *Nature* **222**, 1135–1138 (1969).

51. Vidal, C. M. et al. Age of the oldest known *Homo sapiens* from eastern Africa. *Nature* **601**, 579–583 (2022).

52. Richter, D. et al. The age of the hominin fossils from Jebel Irhoud, Morocco, and the origins of the Middle Stone Age. *Nature* **546**, 293–296 (2017).

53. Berger, L. R. et al. *Homo naledi,* a new species of the genus *Homo* from the Dinaledi Chamber, South Africa.*eLife* **4**, e09560 (2015).

54. Dirks, P. H. et al.The age of *Homo naledi* and associated sediments in the Rising Star Cave, South Africa.*eLife* **6**, e24231 (2017).

55. Kelleher, J., Etheridge, A. M. & McVean, G. Efficient coalescent simulation and genealogical analysis for large sample sizes. *PLoS Comput. Biol.* **12**, e1004842 (2016).

56. Baumdicker, F. et al. Efficient ancestry and mutation simulation with msprime 1.0. *Genetics* **220**, iyab229 (2022).

Acknowledgements

We thank participants for the DNA contributions that enabled this study; in particular, we wish to highlight the generous participation of the Richtersveld Nama community in South Africa and help from local research assistants W. De Klerk and H. Kaimann. Additional assistance and community engagement was conducted by J. Myrick, C. Gignoux, C. Uren and C. Werely. We thank the African Genome Diversity Project for data generation, including T. Carensten, D. Gurdasani and M. Sandhu; L. Anderson-Trocmé and G. Femerling for assistance in creating the map in Fig. 2 and Supplementary Fig. 2, respectively; and N. M. Morales-Garcia for data visualization discussion and designing Figs. 1 and 3. This research was supported by CIHR project grant 437576, Natural Sciences and Engineering Research Council of Canada (NSERC) grant RGPIN-2017-04816, the Canada Research Chair program to S.G. and the Canada Foundation for Innovation, and an NIH grant R35GM133531 to B.M.H.; and E.G.A. was supported by NIH K01 MH121659 and K12 GM102778. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. M.M. and E.H. acknowledge the support of the DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, the South African Medical Research Council Centre for Tuberculosis Research, and the Division of Molecular Biology and Human Genetics at Stellenbosch University, Cape Town, South Africa.

Author information

Authors and Affiliations

1. Department of Integrative Biology, University of Wisconsin–Madison, Madison, WI, USA

   Aaron P. Ragsdale

2. Department of Anthropology, University of California, Davis, CA, USA

   Timothy D. Weaver & Brenna M. Henn

3. Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

   Elizabeth G. Atkinson

4. DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, Stellenbosch University, Cape Town, South Africa

   Eileen G. Hoal & Marlo Möller

5. South African Medical Research Council Centre for Tuberculosis Research, Stellenbosch University, Cape Town, South Africa

   Eileen G. Hoal & Marlo Möller

6. Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

   Eileen G. Hoal & Marlo Möller

7. Genome Center, University of California, Davis, CA, USA

   Brenna M. Henn

8. Department of Human Genetics, McGill University, Montreal, Quebec, Canada

Simon Gravel

Contributions

A.P.R., B.M.H. and S.G. designed the study. B.M.H., E.H. and M.M. designed recruitment protocols and recruited participants. B.M.H. and E.G.A. performed data quality control. A.P.R., B.M.H. and S.G. designed the statistical analyses. A.P.R. conducted the statistical analyses. A.P.R., T.D.W., B.M.H. and S.G. interpreted the results and wrote the first draft of the article. All authors read and edited the paper.

Corresponding authors

Correspondence to Brenna M. Henn or Simon Gravel.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Marta Mirazon Lahr and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information

Supplementary Information

This file contains further discussion, methods and data, Supplementary Tables S1–S8, Supplementary Figs. S1–S40, and Supplementary References.

Reporting Summary

Peer Review File

Supplementary Data

This zipped file contains inferred demographic models in Demes format for all models presented in the main text and all alternative models used in validation of our main results that are discussed in the Supplementary Information.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Ragsdale, A.P., Weaver, T.D., Atkinson, E.G. *et al.* A weakly structured stem for human origins in Africa. *Nature* **617**, 755–763 (2023). https://doi.org/10.1038/s41586-023-06055-y

- Received: 15 March 2022
- Accepted: 05 April 2023
- Published: 17 May 2023
- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-06055-y

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

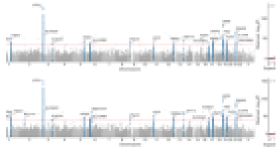
Provided by the Springer Nature SharedIt content-sharing initiative

One species, many roots?

- Eleanor M. L. Scerri

Nature Ecology & Evolution News & Views 17 May 2023

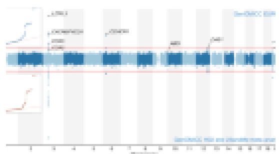
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06055-y



Whole-genome sequencing reveals host factors underlying critical COVID-19

07 March 2022

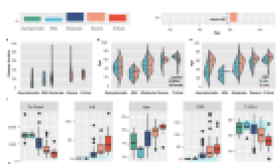
Athanasios Kousathanas, Erola Pairo-Castineira, … J. Kenneth Baillie



Genetic mechanisms of critical illness in COVID-19

11 December 2020

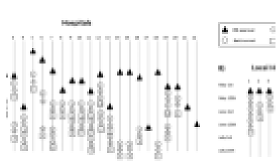
Erola Pairo-Castineira, Sara Clohisey, … J. Kenneth Baillie



Initial whole-genome sequencing and analysis of the host genetic contribution to COVID-19 severity and susceptibility

10 November 2020

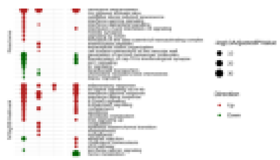
Fang Wang, Shujia Huang, … Lei Liu



[Employing a systematic approach to biobanking and analyzing clinical and genetic data for advancing COVID-19 research]

17 January 2021

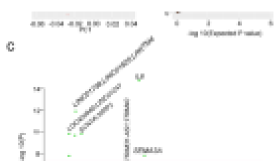
Sergio Daga, Chiara Fallerini, … Elisa Frullanti



[Predicting severity in COVID-19 disease using sepsis blood gene expression signatures]

23 January 2023

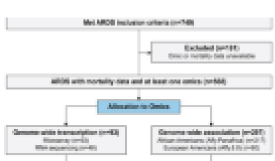
Arjun Baghela, Andy An, … Robert E. W. Hancock



[A genetic variant in IL-6 lowering its expression is protective for critical patients with COVID-19]

02 April 2022

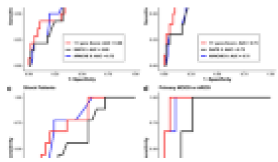
Bo Gong, Lulin Huang, … Zhenglin Yang



[Identification of early and intermediate biomarkers for ARDS mortality by multi-omic approaches]

23 September 2021

S. Y. Liao, N. G. Casanova, … Joe G. N. Garcia



[Prospective validation of an 11-gene mRNA host response score for mortality risk stratification in the intensive care unit]

22 June 2021

Andrew R. Moore, Jonasel Roque, … Angela J. Rogers

[The role of vitamin C in pneumonia and COVID-19 infection in adults with European ancestry: a Mendelian randomisation study]

30 August 2021

L. L. Hui, E. A. S. Nelson, … J. V. Zhao

- Article
- Open Access
- Published: 17 May 2023

GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19

- Erola Pairo-Castineira,
- Konrad Rawlik,

- Andrew D. Bretherick,
- Ting Qi,
- Yang Wu,
- Isar Nassiri,
- Glenn A. McConkey,
- Marie Zechner,
- Lucija Klaric,
- Fiona Griffiths,
- Wilna Oosthuyzen,
- Athanasios Kousathanas,
- Anne Richmond,
- Jonathan Millar,
- Clark D. Russell,
- Tomas Malinauskas,
- Ryan Thwaites,
- Kirstie Morrice,
- Sean Keating,
- David Maslove,
- Alistair Nichol,
- Malcolm G. Semple,
- Julian Knight,
- Manu Shankar-Hari,
- Charlotte Summers,
- Charles Hinds,
- Peter Horby,
- Lowell Ling,
- Danny McAuley,
- Hugh Montgomery,
- Peter J. M. Openshaw,
- Colin Begg,
- Timothy Walsh,
- Albert Tenesa,
- Carlos Flores,
- José A. Riancho,
- Augusto Rojas-Martinez,
- Pablo Lapunzina,
- GenOMICC Investigators,

- SCOURGE Consortium,
- ISARICC Investigators,
- The 23andMe COVID-19 Team,
- Jian Yang,
- Chris P. Ponting,
- James F. Wilson,
- Veronique Vitart,
- Malak Abedalthagafi,
- Andre D. Luchessi,
- Esteban J. Parra,
- Raquel Cruz,
- Angel Carracedo,
- Angie Fawkes,
- Lee Murphy,
- Kathy Rowan,
- Alexandre C. Pereira,
- Andy Law,
- Benjamin Fairfax,
- Sara Clohisey Hendry &
- …
- J. Kenneth Baillie

Show authors

Nature **volume 617**, pages 764–768 (2023)

- 21k Accesses
- 617 Altmetric
- Metrics details

Abstract

Critical illness in COVID-19 is an extreme and clinically homogeneous disease phenotype that we have previously shown[1] to be highly efficient for discovery of genetic associations[2]. Despite the advanced stage of illness at

presentation, we have shown that host genetics in patients who are critically ill with COVID-19 can identify immunomodulatory therapies with strong beneficial effects in this group[3]. Here we analyse 24,202 cases of COVID-19 with critical illness comprising a combination of microarray genotype and whole-genome sequencing data from cases of critical illness in the international GenOMICC (11,440 cases) study, combined with other studies recruiting hospitalized patients with a strong focus on severe and critical disease: ISARIC4C (676 cases) and the SCOURGE consortium (5,934 cases). To put these results in the context of existing work, we conduct a meta-analysis of the new GenOMICC genome-wide association study (GWAS) results with previously published data. We find 49 genome-wide significant associations, of which 16 have not been reported previously. To investigate the therapeutic implications of these findings, we infer the structural consequences of protein-coding variants, and combine our GWAS results with gene expression data using a monocyte transcriptome-wide association study (TWAS) model, as well as gene and protein expression using Mendelian randomization. We identify potentially druggable targets in multiple systems, including inflammatory signalling (*JAK1*), monocyte–macrophage activation and endothelial permeability (*PDE4A*), immunometabolism (*SLC2A5* and *AK5*), and host factors required for viral entry and replication (*TMPRSS2* and *RAB2A*).

Main

The design of the GenOMICC study and the rationale for focusing on critical illness has been previously described[1,2]. In brief, patients with confirmed COVID-19 requiring continuous cardiorespiratory monitoring or organ support (a generalizable definition for critical illness) were recruited in 2020–2022. We first performed ancestry-specific GWAS analyses according to the methods that we described previously[1,2]. Using the results of these GWAS analyses, previously reported results obtained using GenOMICC participants with whole-genome sequencing data[2] and data from GenOMICC Brazil, we performed trans-ancestry and -platform meta-analyses within the GenOMICC study for a critically ill COVID-19 phenotype and a hospitalized COVID-19 phenotype (Extended Data Fig. 1). The results of these GenOMICC-only meta-analyses are presented for both critically ill and hospitalized phenotypes (Table 1 and Extended Data Fig. 2).

To put these results into the context of existing knowledge, we performed comprehensive meta-analyses, drawing on further GWAS results, including data shared by the SCOURGE consortium and published data from the COVID-19 Human Genetics Initiative (HGIv6, 2021)[4]. The characteristics of the contributing studies are summarized in Supplementary Tables 13 and 14 for the critically ill and hospitalized phenotypes, with further details on each study provided in the Supplementary Information. We used a mathematical subtraction approach, as done in our previous work[2], to remove signals of previous GenOMICC releases from HGIv6, yielding an independent dataset.

Table 1 Genome-wide significant associations with critical COVID-19, listing independent lead variants

As no replication cohorts exist for these meta-analyses, we used the heterogeneity across studies to assess the reliability of individual findings (Supplementary Table 15). Owing to the unusually extreme phenotype in the GenOMICC study, some heterogeneity is expected for the strongest associations when compared with studies with more permissive inclusion criteria. Importantly, significant heterogeneity was not detected for any of the findings that we report here (Supplementary Table 15). Comparing effect estimates between studies using a regression approach that takes into account estimation errors (Methods), we detected systematic differences in effect sizes between studies (Extended Data Fig. 3). For example, effects for the HGI critical illness phenotype (which was designed to parallel the GenOMICC inclusion criteria) are smaller than those obtained using prospective recruitment in GenOMICC by a factor of 0.68. As the effect sizes in GenOMICC are consistently larger than other studies, and GenOMICC contributes a disproportionately large signal to meta-analyses of both critical and hospitalized phenotypes (Extended Data Fig. 4), between-study heterogeneity is likely to reflect the careful case ascertainment and extreme phenotype in GenOMICC compared with other studies.

We found 49 common genetic associations with critical COVID-19 meeting our criteria for genome-wide significance in the absence of heterogeneity (Extended Data Fig. 2 and Table 1). Findings from previous reports were consistently replicated (Extended Data Table 2). Conditional analysis revealed two additional lead variants (Table 1) and statistical fine-mapping

provided credible sets of putative causal variants for a majority of lead variants (Supplementary Figs. 27–44 and Supplementary Table 5). Gene-level analyses found 196 significantly associated genes at a Bonferroni-corrected threshold (Supplementary Table 10). There were no genome-wide significant differences in the effects between sexes in a sex-stratified meta-analysis using a subset of cohorts (Supplementary Fig. 1).

Therapeutic implications

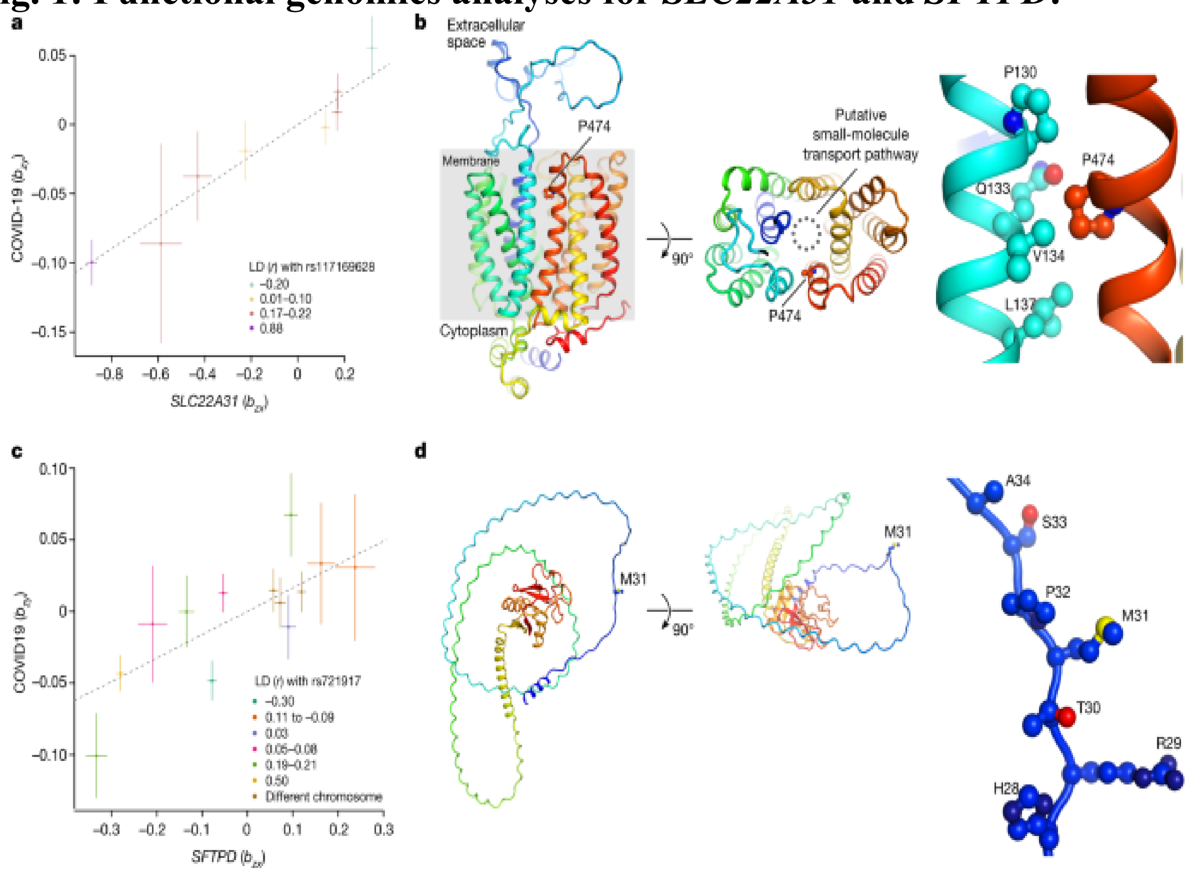
Our analysis is limited to common variants that are detectable on genotyping arrays and imputation panels. Although most lead variants are not directly causal, in some cases, they highlight molecular mechanisms that alter clinical outcomes in COVID-19, and may have direct therapeutic relevance. To investigate the disease mechanisms, we first quantified the effect of inferred gene expression on critical illness in three relevant tissue/cell types. Many of the genes that we have found to be implicated in critical COVID-19 (refs. [1,2]) are highly expressed in the monocyte–macrophage system, which has poor coverage in existing expression quantitative trait loci (eQTL) datasets. For this reason, we constructed a new TWAS model in primary monocytes obtained from 176 individuals (Methods). We found significant associations after Bonferroni correction between critical COVID-19 and predicted gene expression in lung (33), blood (21), monocyte (37) and all-tissue (107) meta-analysis (Supplementary Table 2 and Supplementary Table 11). We extended these findings using generalized summary-level data Mendelian randomization (GSMR) for RNA expression (Fig. 2, Extended Data Table 1, Supplementary Figs. 11–18 and Supplementary Table 4).

In parallel, we assessed the effect of genetically determined variation in circulating protein levels on the critical illness phenotype using GSMR[5]. We identified 15 unique proteins linked to critical illness, as summarized in Extended Data Table 1 (Supplementary Table 3). Of the significant results, we found causal evidence implicating five new proteins in comparison to our previous GSMR analysis[2]: QSOX2, CREB3L4, myeloperoxidase (MPO), ADAMTS13 and mannose-binding lectin-2 (MBL2) (Supplementary Fig. 10). These include well-studied biomarkers and potential drug targets in sepsis—the innate immune pattern recognition receptor MBL2 and the neutrophil effector enzyme MPO. ADAMTS13

modulates von Willebrand-factor-mediated platelet thrombus formation and may have a role in the hypercoagulable state in critical COVID-19 (Extended Data Fig. 5).

Three genes containing non-synonymous protein-coding changes associated with severe disease were also found to have significant effects from differential gene expression: *SLC22A31* (ref. [2]) (Fig. 1), *SFTPD*[4] (Fig. 1) and *TKY2* (ref. [1]) (Extended Data Fig. 6). Further biological and clinical research will be required to dissect the genetic evidence at these loci. In the example of *TYK2*, there is now a therapeutic test of the genetic predictions. Our previous report of association between higher expression and critical illness[1] led directly to the inclusion of a new drug, baricitinib, in a large clinical trial; the result demonstrated a clear therapeutic benefit[3]. This therapeutic signal is consistent across multiple trials, providing the first proof-of-concept for drug target identification using genetics in critical illness and infectious disease.

Fig. 1: Functional genomics analyses for *SLC22A31* and *SFTPD*.

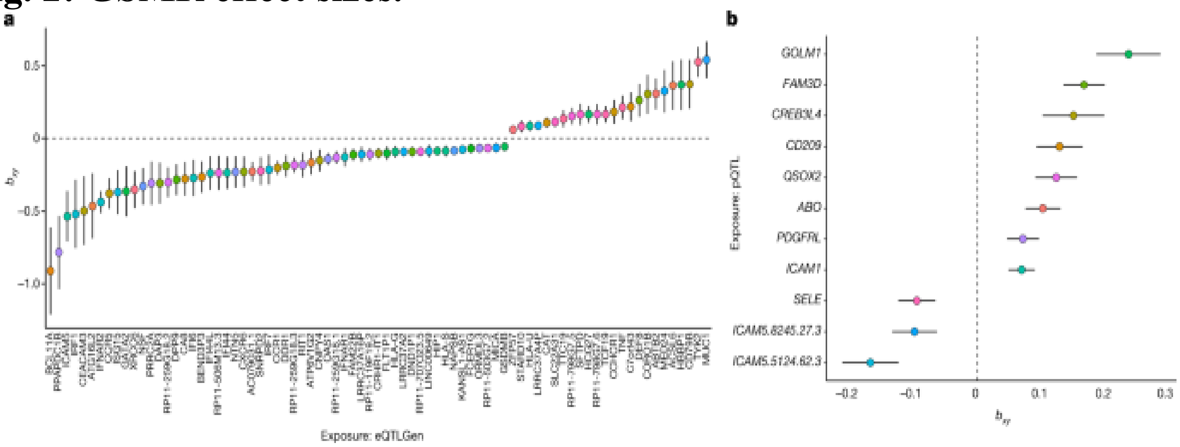


a, Effect-size plot for the effect of multiple variants on *SLC22A31* expression (eQTLgen, *x* axis) against increasing susceptibility to critical COVID-19 ($\beta_{xy} = 0.11$; $P_{xy} = 1.3 \times 10^{-9}$). The colour shows linkage disequilibrium (LD) with the missense variant rs117169628. **b**, Three cartoon views of an AlphaFold[22] model of putative solute carrier family 22 member 31 (SLC22A31; UniProtKB: A6NKX4). The side chains of Pro474 and interacting amino acids are shown as connected spheres. A putative channel for small-molecule transport across the cell membrane is indicated by a dashed circle. Pro474 is predicted to be located in the transmembrane helix and point towards a putative transport pathway of a small molecule. The risk variant, P474L (Ala at rs117169628) would be expected to introduce more flexibility to the transmembrane helix and might therefore affect the transport properties of SLC22A31. Pro474 is predicted to be in a tightly packed environment, and may therefore affect the folding of SLC22A31. **c**, Effect-size plot for effect of multiple variants on SFTPD expression (eQTLgen, *x* axis) against increasing susceptibility to critical COVID-19 ($\beta_{xy} = 0.16$; $P_{xy} = 9.7 \times 10^{-6}$). Colour shows linkage disequilibrium with the missense variant rs721917. **d**, Three cartoon views of an AlphaFold[22] model of pulmonary surfactant-associated protein D (SFTPD; UniProtKB: P35247). The side chain of the variant Met31 is shown as connected spheres. Met31 is predicted to be located in the secondary-structure-lacking region of SFTPD. In the diagram on the right, oxygen and nitrogen atoms are coloured red and blue respectively, and the sulfur atom is coloured yellow.

To assess the immediate therapeutic use of our results for repurposing of existing compounds, we considered the drug therapies under consideration by the UK COVID-19 Therapeutic Advisory Panel (UK-CTAP), a national independent review group supported by an expert due-diligence panel[6]. Consistent evidence from gene-level GWAS (Supplementary Table 6 and Supplementary Table 10) and post-GWAS analyses was identified for several licensed compounds (Supplementary Table 12). For example, we found an association in another gene encoding a protein that is inhibited by baricitinib and other JAK inhibitors—the intracellular signalling kinase, *JAK1*, which is stimulated by numerous cytokines including type I interferons and IL-6. Mendelian randomization analysis of RNA expression revealed a significant positive association between the expression of the

gene encoding a canonical inflammatory cytokine, tumour necrosis factor (*TNF*), and severe disease (Fig. 2). This suggests that inhibition of TNF signalling may be an effective therapy in severe COVID-19.

Fig. 2: GSMR effect sizes.



a,**b**, The predicted effect of change in protein concentration (**a**) and gene expression (**b**) on the risk of critical COVID-19 is shown for proteins and genes significantly linked to critical COVID-19 by GSMR (false-discovery rate (FDR) < 0.01). The bars show 95% confidence intervals.

Our additional expression data in monocytes reveal a marked tissue-specific effect on expression of *PDE4A*. This phosphodiesterase regulates the production of multiple inflammatory cytokines by myeloid cells. In contrast to the negative correlations seen in the lungs and blood, we show that a genetic tendency for higher expression of *PDE4A* in monocytes is associated with critical COVID-19 (Supplementary Table 11). Inhibition of PDE4A by several existing drugs is under investigation in multiple inflammatory diseases[7], reduces pulmonary endothelial permeability[8] and appears to be safe in small clinical trials in patients with COVID-19.

The postulated biological role of genes associated with critical COVID-19 in GWAS, TWAS and GSMR results is shown in Extended Data Fig. 5, which highlights the preponderance of genes with expression or functions in the mononuclear phagocyte system. This includes *SLC2A5*, encoding the GLUT5 fructose transporter, which is strongly inducible in primary macrophages in response to inflammatory stimulation[9], and *XCR1*, a dendritic cell receptor with a critical role in cytotoxic T cell-mediated

antiviral immunity[10]. *NPNT*, a significant meta-TWAS association in the genome-wide significant region on chromosome 4 (chr4:105673359; Supplementary Table 11), encodes a pulmonary basement membrane protein that may have a protective role in acute lung injury[11].

Host–pathogen interaction

Our results also demonstrate the capacity of host genetics to reveal core mechanisms of disease. Multiple genes implicated in viral entry are associated with severe disease. In addition to *ACE2*, we detect a genome-wide significant association in *TMPRSS2*, a key host protease that facilitates viral entry that we have previously studied as a candidate gene[12]. This effect may be viral-lineage specific[13]. A strong GWAS association is seen in *RAB2A* (Table 1), with TWAS evidence suggesting that more expression of this gene is associated with worse disease (Supplementary Table 11). *RAB2A* is highly ranked in our previous meta-analysis by information content[14] study of host genes implicated in SARS-CoV-2 interaction using in vitro and clinical data[15], and is consistent with CRISPR screen data showing that *RAB2A* is required for viral replication[16].

Although our focus on critical illness enhances discovery power (Extended Data Fig. 4), it has the disadvantage of combining genetic signals for multiple stages in disease progression, including viral exposure, infection and replication, and development of inflammatory lung disease. From these data alone we cannot identify when in disease progression the causal effect is mediated, although clinical evidence helps to make some predictions[17] (Extended Data Fig. 5). As most cases included were recruited before vaccinations and treatments became available (Extended Data Fig. 7), at present, our study does not have sufficient statistical power to dissect the genetic effects of treatments or vaccination. These effects may include the masking of true associations, or the detection of genetic effects mediated by vaccine or drug response, rather than COVID-19 susceptibility. However, the absence of divergent genetic effects between studies (Supplementary Figs. 2–5) or consistent changes in effect allele frequency among cases over time (Supplementary Figs. 45–48) suggests that treatment and vaccination

have not substantially affected the association between the specific variants that we report and the risk of critical illness.

As we performed a meta-analysis of multiple studies that may have slightly different definitions of the phenotype, effect sizes differ between studies (Supplementary Figs. 2–5). This, together with ancestry-specific effects[1], may explain the heterogeneity in strong GWAS signals, such as the LZTFL1 signal in Table 1. Different studies also have sets of variants that are not completely overlapping, so *P* values between variants in high linkage disequilibrium are more different than expected. Although most of the studies contain individuals from multiple ancestries, a large majority of the individuals are of European ancestry. In future research, there is a scientific and moral imperative to include the full diversity of human populations.

Together, these results deepen our understanding of the pathogenesis of critical COVID-19 and highlight new biological mechanisms of disease, several of which have immediate potential for therapeutic targeting.

Methods

Hospitalization meta-analysis

The hospitalized phenotype includes patients who were hospitalized with a laboratory-confirmed SARS-Cov2 infection. In this analysis we included GenOMICC, GenOMICC Brazil, GenOMICC Saudi Arabia, ISARIC4C, HGIv6 B2 phenotype with subtraction of GenOMICC data, SCOURGE hospitalized versus population and mild cases, and 23andMe broad respiratory phenotype. A summary description of each analysis is given above, a table with the included studies can be found in Supplementary Table 14 and an extended description can be found in Supplementary Table 1.

Critical illness meta-analysis

The critically ill COVID-19 group included patients who were hospitalized owing to symptoms associated with laboratory-confirmed SARS-CoV-2 infection and who required respiratory support or whose cause of death was

associated with COVID-19. In the critical illness analysis, we included GenOMICC, patients with critical illness from ISARIC4C, HGIv6 phenotype A2 with subtraction of GenOMICC data, SCOURGE severity grades 3 and 4 versus population controls, and 23andMe respiratory support phenotype. A summary description of each analysis can be found above, a table with the included studies can be found in Supplementary Table 13 and an extended description can be found in Supplementary Table 1.

Meta-analyses

All meta-analyses across studies were performed using a fixed-effect inverse-variance weighting method and control for population stratification in the METAL software[23]. Allele frequency was calculated as the average frequency across studies with the METAL option AVERAGEFREQ. *P* values for heterogeneity in effect sizes between studies were calculated using a Cochran's *Q*-test implemented in METAL. For variants in the same position with different REF and ALT alleles across studies, the GenoMICC variant in the European population was selected and the rest were removed. Finally, variants with switched ALT and REF alleles between HGIv6 and GenOMICC were also removed on the basis of differences in allele frequency of the alternative allele. Variants were annotated to the closest genes using dbsnp v.b151 GRCh38p7 and bionrRt R package (v.2.46.3)[24]. As each single-nucleotide polymorphism (SNP) of the meta-analysis can be present in different subsets of cohorts, there may be large differences in *P* values in SNPs with a high level of linkage disequilibrium, which may have an effect on downstream analyses. For this reason, variants that were not present in one of the three biggest studies—GenOMICC European ancestry, HGIv6 or SCOURGE—were filtered out from post-GWAS analysis.

Conditional analysis

We performed a step-wise conditional analysis to find independent signals. As European-specific data are not available in some cohorts but European ancestry is largely predominant (87.2% of cases with critical illness), we performed the conditional analysis using a European reference panel and the meta-analysis results of the whole cohort. To perform the conditional analysis, we used the GCTA (v.1.9.3) --cojo-slct function[25]. The parameters

for the function were $P = 5 \times 10^{-8}$, a distance of 10,000 kb and a co-linear threshold of 0.9 (ref. [26]), and the reference population for the conditional analysis was individuals of European ancestry with whole-genome sequence available in the GenOMICC study and whole genomes from the 100,000 Genomics England project[2].

Credible set fine-mapping

We performed fine-mapping using the SuSiE model[27] to construct credible sets for the independent signals identified using conditional analysis. As for conditional analysis, we used a European reference panel and the meta-analysis results of the whole critical illness cohort. We performed analyses in 1 Mb windows centred on the lead variants identified through conditional analysis. In cases in which windows for multiple variants overlapped, they were joined into a single window. For each window, we fitted the SuSiE summary statistics model setting the expected number of independent signals to the number of identified though conditional analysis. Models for three windows did not converge in 500 iterations and have been excluded. As a reference, we used the publically available linkage disequilibrium information for non-Finish Europeans from the GNOMAD 2.1.1 release. Full data for all variants included in credible sets are included in Supplementary Table 5.

Gene-level analysis

We performed an analysis summarizing the genetic associations at the gene level using the mBAT-combo method[28]. We used the COVID 'all critical cohorts' meta-analysis (GenOMICC, HGIv6 phenotype A2, SCOURGE and 23andMe) summary statistics. As this is a trans-ethnic meta-analysis, we used a mixed ancestry linkage disequilibrium reference panel, consisting of 3,202 1000 Genomes phase 3 samples. We considered a list of protein-coding genes with unique ensemble gene ID based on the release from GENCODE (v.40) for hg38, which can be found on the mBAT-combo website (https://yanglab.westlake.edu.cn/software/gcta/#mBAT-combo). A gene region was taken to span 50 kb upstream to 50 kb downstream of the gene's untranslated regions.

Sex-stratified meta-analysis

To test for differences in genetic effects, we performed sex-stratified GWAS of the COVID-19 critical illness phenotype in the European ancestry GenOMICC WGS and genotyped cohorts and SCOURGE. We then performed a meta-analysis for each sex following the same methods as for the main analysis. We tested for differences in effects between the meta-analyses of the two sexes following previously described methods[29].

Mendelian randomization

GSMR[5] was performed. We used the COVID 'all critical cohorts' meta-analysis (GenOMICC, HGIv6 phenotype A2, SCOURGE and 23andMe) as the outcome, protein expression quantitative-trait loci (pQTLs) from ref. [30] and RNA expression quantitative-trait loci (eQTLs) from eQTLgen[31] (2019-12-23 data release) as exposures, and 10,000 individuals of European ancestry randomly sampled from the UK Biobank as the linkage disequilibrium reference cohort (50,000 for linkage disequilibrium to missense variant plots). GSMR was performed for all exposures for which we were able to identify two or more suitable SNPs. SNPs were chosen to meet the following criteria: (1) SNP to exposure association $P < 5 \times 10^{-8}$; (2) linkage disequilibrium clumping lead SNPs only (±1 Mb, $r^2 < 0.05$); (3) SNP not removed by HEIDI-outlier filtering (for the removal of SNPs with evidence of horizontal pleiotropy) at the default threshold value of 0.01. eQTLGen effect sizes and standard errors were estimated as described in supplementary note 2 of ref. [32]. We considered as significant those exposure–outcome pairs with FDR < 0.05.

TWAS analysis

To perform TWAS analysis in GTExv8 tissues[33], we used the MetaXcan framework and the GTExv8 eQTL and sQTL MASHR-M models available for download online (http://predictdb.org/) and the 'all critical cohorts' meta-analysis. We first calculated individual TWAS for whole blood and lungs using the S-PrediXcan function[34,35]. We next performed a metaTWAS including data from all tissues to increase the statistical power using s-

MultiXcan[36]. We applied Bonferroni correction to the results to choose significant genes and introns for each analysis.

Monocyte gene expression

To detect eQTLs, untreated primary monocytes were prepared from 174 healthy individuals of Northern European (British) ancestry recruited through the Oxford Biobank. Poly(A) RNA was paired-end 100 bp sequenced in the Oxford Genome Centre using the Illumina HiSeq-4000 machines (median = 47,735,438 reads per sample). Reads were aligned to CRGh38/hg38 using HISAT2 with the default parameters. High mapping quality reads were selected on the basis of MAPQ score using bamtools. Duplicate reads were marked and removed using picard (v.1.105). Samtools was used to pass through the mapped reads and calculate statistics. Read count information was generated using HTSeq and normalized using DESeq2. Sample contamination and swaps were detected by comparing the imputed SNP-array genotypes with genotypes called from RNA-seq using verifyBamID. Genotyping was performed with Illumina HumanOmniExpress with coverage of 733,202 separate markers. Genotypes were pre-phased with SHAPEIT2, and missing genotypes were imputed with PBWT. Poly(A) RNA was paired-end sequenced at the Oxford Genome Centre using the Illumina HiSeq-4000 machines. vcftools (v0.1.12b) was applied on genetic variation data in the form of variant call format (VCF) files to filter out indels and SNPs with a minor allele frequency of less than 0.04.

TWAS analysis for monocyte data was performed using genotyping and monocyte RNA-sequencing data from 174 individuals. Using a region of 500 kb around each gene, we calculated gene expression models using the Fusion R package[37]. For each gene, three models were calculated adding as covariates the two first principal components calculated from the genotype: blup, elastic networks and lasso. The model with a better r^2 between predicted and measured expression in a fivefold cross-validation was chosen. Then SNP genetic heritability was calculated for the 500 kb region for each gene and those genes with a nominal significant SNP heritability estimate ($P \leq 0.01$) were chosen for the TWAS analysis. Summary statistics

for the ‘all critical cohorts’ meta-analysis and the best model for each gene were then used to perform the TWAS.

Colocalization

Significant genes in the TWAS and metaTWAS were selected for a colocalization analysis using the coloc R package. The lead SNPs and a region of 200 Mb around the gene were used to colocalize with significant genes in the TWAS with eQTL summary statistics data on the region from GTExv8 lung, GTExv8 whole blood, eQTLgen or monocyte eqtl. As in our previous analysis[2], we first performed a sensitivity analysis of the posterior probability of colocalization (PPH4) on the prior probability of colocalization (P12), going from $P12 = 10^{-8}$ to $P12 = 10^{-4}$, with the default threshold being $P12 = 10^{-5}$. eQTL signal and GWAS signals were deemed to colocalize if these two criteria were met: (1) at $P12 = 5 \times 10^{-5}$ the probability of colocalization $PPH4 > 0.5$; and (2) at $P12 = 10^{-5}$ the probability of independent signal (PPH3) was not the main hypothesis ($PPH3 < 0.5$). These criteria were chosen to allow eQTLs with weaker *P* values, owing to lack of power in GTEx v.8, to be colocalized with the signal when the main hypothesis using small priors was that there was not any signal in the eQTL data.

Effect comparison

We compared the estimates of effect sizes between the individual GWASs used in the meta-analysis, for all variants that were genome-wide significant in at least one of the individual GWASs. To this end, we regressed the effects obtained using critical illness and hospitalization in the SCOURGE and 23andMe cohorts, as well as the HGI meta-analyses on the effect estimates obtained using the GenOMICC cohort. To account for estimation errors present in both the dependent and independent variables of the regression we used orthogonal distance regression[38].

Weight of studies

To calculate the weight of GenOMICC, we downloaded the leave-one-out data of HGIv7. As the meta-analysis is performed using a variance-weighted

method, we can recover the variance for each SNP as $v=\frac{1}{{{\rm{s.e.}}}^{2}}$, for the meta-analysis of all of the cohorts and for each one of the leave-one-out analysis. The total weight is ${w}_{{\rm{tot}}}=\frac{1}{v}$ and the weight leaving out a specific study is ${w}_{{\rm{loo}}}=\frac{1}{{v}_{{\rm{loo}}}}$. The weight of a cohort is then ${w}_{{\rm{tot}}}-{w}_{{\rm{loo}}}$. We calculated the weight for each the significant SNPs in our analysis for each study and normalized it using the total weight. Finally, we calculated the mean and s.d. from the significant SNPs for each cohort.

Forest plots

To compare effects between cohorts, we first performed a trans-ancestry meta-analysis for GenOMICC and 23andMe using METAL[23]. Then, we used the metagen and forest functions of the meta R package to produce forest plots for critical illness and hospitalization separately.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Downloadable summary data are available through the GenOMICC data site (https://genomicc.org/data). Summary statistics are available, but without the 23andMe summary statistics, except for the 10,000 most significant hits, for which full summary statistics are available. The full GWAS summary statistics for the 23andMe discovery dataset will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. For further information and to apply for access to the data, see the 23andMe website (https://research.23andMe.com/dataset-access/). All individual-level genotype and whole-genome sequencing data (for both academic and commercial uses) can be accessed through the UKRI/HDR UK Outbreak Data Analysis Platform (https://odap.ac.uk). A restricted dataset for a subset of GenOMICC participants is also available through the Genomics England

data service. Monocyte RNA-seq data are available under the title 'Monocyte gene expression data' within the Oxford University Research Archives (https://doi.org/10.5287/ora-ko7q2nq66). Sequencing data will be made freely available to organizations and researchers to conduct research in accordance with the UK Policy Framework for Health and Social Care Research through a data access agreement. Sequencing data have been deposited at the European Genome–Phenome Archive (EGA), which is hosted by the EBI and the CRG, under accession number EGAS00001007111.

Code availability

Code to calculate the imputation of *P* values on the basis of SNPs in linkage disequilibrium is available at GitHub (https://github.com/baillielab/GenOMICC_GWAS).

References

1. Pairo-Castineira, E. et al. Genetic mechanisms of critical illness in COVID-19. *Nature* **591**, 92–98 (2021).

2. Kousathanas, A. et al. Whole-genome sequencing reveals host factors underlying critical COVID-19. *Nature* **607**, 97–103 (2022).

3. Abani, O. et al. Baricitinib in patients admitted to hospital with COVID-19 (RECOVERY): a randomised, controlled, open-label, platform trial and updated meta-analysis. *Lancet* **400**, 359–368 (2022).

4. Pathak, G. A. et al. A first update on mapping the human genetic architecture of COVID-19. *Nature* **608**, E1–E10 (2022).

5. Zhu, Z. et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).

6. Chinnery, P. F. et al. Choosing drugs for UK COVID-19 treatment trials. *Nat. Rev. Drug Discov.* **21**, 81–82 (2022).

7. Peng, T., Qi, B., He, J., Ke, H. & Shi, J. Advances in the development of phosphodiesterase-4 inhibitors. *J. Med. Chem.* **63**, 10594–10617 (2020).

8. Sanz, M.-J. et al. Roflumilast inhibits leukocyte-endothelial cell interactions, expression of adhesion molecules and microvascular permeability. *Br. J. Pharmacol.* **152**, 481–92 (2007).

9. Forrest, A. R. R. et al. A promoter-level mammalian expression atlas. *Nature* **507**, 462–470 (2014).

10. Brewitz, A. et al. CD8. *Immunity* **46**, 205–219 (2017).

11. Wilson, C. L., Hung, C. F. & Schnapp, L. M. Endotoxin-induced acute lung injury in mice with postnatal deletion of nephronectin. *PLoS ONE* **17**, e0268398 (2022).

12. David, A. et al. A common TMPRSS2 variant has a protective effect against severe COVID-19. *Curr. Res. Transl. Med.* **70**, 103333 (2022).

13. Meng, B. et al. Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts tropism and fusogenicity. *Nature* https://doi.org/10.1038/s41586-022-04474-x (2022).

14. Li, B. et al. Genome-wide CRISPR screen identifies host dependency factors for influenza A virus infection. *Nat. Commun.* **11**, 164 (2020).

15. Parkinson, N. et al. Dynamic data-driven meta-analysis for prioritisation of host genes implicated in COVID-19. *Sci. Rep.* **10**, 22303 (2020).

16. Hoffmann, H.-H. et al. Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors. Preprint at *bioRxiv* https://doi.org/10.1101/2020.09.11.291716 (2020).

17. Russell, C. D., Lone, N. I. & Baillie, J. K. Comorbidities, multimorbidity and COVID-19. *Nat. Med.* **29**, 334–343 (2023).

18. Niemi, M. E. K. et al. Mapping the human genetic architecture of COVID-19. *Nature* **600**, 472–477 (2021).

19. Ellinghaus, D. et al. Genomewide association study of severe COVID-19 with respiratory failure. *New Engl. J. Med.* **383**, 1522–1534 (2020).

20. Cruz, R. et al. Novel genes and sex differences in COVID-19 severity. *Hum. Mol. Genet.* **31**, 3789–3806 (2022).

21. Degenhardt, F. et al. Detailed stratified GWAS analysis for severe COVID-19 in four European populations. *Hum. Mol. Genet.* **31**, 3945–3966 (2022).

22. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).

23. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).

24. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of genomic datasets with the r/bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).

25. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).

26. Yang, J. et al. Conditional and joint multiple-SNP analysis of GWAS summary statistics identifies additional variants influencing complex traits. *Nat. Genet.* **44**, 369–375 (2012).

27. Wang, G., Sarkar, A., Carbonetto, P. & Stephens, M. A simple new approach to variable selection in regression, with application to genetic fine mapping. *J. R. Stat. Soc. B* **82**, 1273–1300 (2020).

28. Li, A. et al. mBAT-combo: a more powerful test to detect gene-trait associations from GWAS data. Preprint at *bioRxiv* https://doi.org/10.1101/2022.06.27.497850 (2022).

29. Bernabeu, E. et al. Sex differences in genetic architecture in the UK Biobank. *Nature* **53**, 1283–1289 (2021).

30. Sun, B. B. et al. Genomic atlas of the human plasma proteome. *Nature* **558**, 73–79 (2018).

31. Võsa, U. et al. Large-scale *cis*- and *trans*-eQTL analyses identify thousands of genetic loci and polygenic scores that regulate blood gene expression. *Nat. Genet.* **53**, 1300–1310 (2021).

32. Zhu, Z. et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat. Genet.* **48**, 481–487 (2016).

33. The GTEX Consortium. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science* **369**, 1318–1330 (2020).

34. Gamazon, E. R. et al. A gene-based association method for mapping traits using reference transcriptome data. *Nat. Genet.* **47**, 1091–1098 (2015).

35. Barbeira, A. N. et al. Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. *Nat. Commun.* **9**, 1825 (2018).

36. Barbeira, A. N. et al. Integrating predicted transcriptome from multiple tissues improves association detection. *PLoS Genet.* **15**, e1007889 (2019).

37. Gusev, A. et al. Integrative approaches for large-scale transcriptome-wide association studies. *Nat. Genet.* **48**, 245–252 (2016).

38. Boggs, P. T. & Rogers, J. E. Orthogonal distance regression. *Contemp. Math.* **112**, 183–194 (1990).

39. Liang, J. et al. Lead identification of novel and selective TYK2 inhibitors. *Eur. J. Med. Chem.* **67**, 175–87 (2013).

40. Li, Z. et al. Two rare disease-associated Tyk2 variants are catalytically impaired but signaling competent. *J. Immunol.* **190**, 2335–2344 (2013).

Acknowledgements

We thank the patients and their loved ones who volunteered to contribute to this study at one of the most difficult times in their lives, and the research staff in every intensive care unit who recruited patients at personal risk during the most extreme conditions ever witnessed in most hospitals. GenOMICC was funded by Sepsis Research (the Fiona Elizabeth Agnew Trust), the Intensive Care Society, a Wellcome Trust Senior Research Fellowship (to J.K.B., 223164/Z/21/Z), the Department of Health and Social Care (DHSC), Illumina, LifeArc, the Medical Research Council, UKRI, a BBSRC Institute Program Support Grant to the Roslin Institute (BBS/E/D/20002172, BBS/E/D/10002070 and BBS/E/D/30002275) and UKRI grants MC_PC_20004, MC_PC_19025, MC_PC_1905 and MRNO2995X/1. A.D.B. acknowledges funding from the Wellcome PhD training fellowship for clinicians (204979/Z/16/Z), the Edinburgh Clinical Academic Track (ECAT) programme. This research is supported in part by the Data and Connectivity National Core Study, led by Health Data Research UK in partnership with the Office for National Statistics and funded by UK Research and Innovation (grant MC_PC_20029). Laboratory work was funded by a Wellcome Intermediate Clinical Fellowship to B.F. (201488/Z/16/Z). We acknowledge the staff at NHS Digital, Public Health England and the Intensive Care National Audit and Research Centre who provided clinical data on the participants; and the National Institute for Healthcare Research Clinical Research Network (NIHR CRN) and the Chief Scientist's Office (Scotland), who facilitate recruitment into research studies in NHS hospitals, and to the global ISARIC and InFACT consortia. GenOMICC genotype controls were obtained using UK Biobank Resource under project 788 funded by Roslin Institute Strategic Programme Grants from the BBSRC (BBS/E/D/10002070 and BBS/E/D/30002275) and Health Data Research UK (HDR-9004 and HDR-9003). UK Biobank data were used in the GSMR analyses presented here under project 66982. The UK Biobank was established by the Wellcome Trust medical charity, Medical Research Council, Department of Health, Scottish Government and the Northwest Regional Development Agency. It has also had funding from the Welsh Assembly Government, British Heart Foundation and Diabetes UK. The work of L.K. was supported by an RCUK Innovation Fellowship from

the National Productivity Investment Fund (MR/R026408/1). J.Y. is supported by the Westlake Education Foundation. SCOURGE is funded by the Instituto de Salud Carlos III (COV20_00622 to A.C., PI20/00876 to C.F.), European Union (ERDF) ‘A way of making Europe’, Fundación Amancio Ortega, Banco de Santander (to A.C.), Cabildo Insular de Tenerife (CGIEU0000219140 ‘Apuestas científicas del ITER para colaborar en la lucha contra la COVID-19’ to C.F.) and Fundación Canaria Instituto de Investigación Sanitaria de Canarias (PIFIISC20/57 to C.F.). We also acknowledge the contribution of the Centro National de Genotipado (CEGEN) and Centro de Supercomputación de Galicia (CESGA) for funding this project by providing supercomputing infrastructures. A.D.L. is a recipient of fellowships from the National Council for Scientific and Technological Development (CNPq)-Brazil (309173/2019-1 and 201527/2020-0). We thank the members of the Banco Nacional de ADN and the GRA@CE cohort group; and the research participants and employees of 23andMe for making this work possible. A full list of contributors who have provided data that were collated in the HGI project, including previous iterations, is available online (https://www.covid19hg.org/acknowledgements).

Author information

Author notes

1. These authors contributed equally: Erola Pairo-Castineira, Konrad Rawlik

Authors and Affiliations

1. Baillie Gifford Pandemic Science Hub, Centre for Inflammation Research, The Queen’s Medical Research Institute, University of Edinburgh, Edinburgh, UK

   Erola Pairo-Castineira, Konrad Rawlik, Andrew D. Bretherick, Marie Zechner, Fiona Griffiths, Wilna Oosthuyzen, Jonathan Millar, Clark D. Russell, Sara Clohisey Hendry & J. Kenneth Baillie

2. MRC Human Genetics Unit, Institute of Genetics and Cancer, University of Edinburgh, Western General Hospital, Edinburgh, UK

Erola Pairo-Castineira, Andrew D. Bretherick, Lucija Klaric, Anne Richmond, Albert Tenesa, Alison M. Meynert, Murray Wham, Chris P. Ponting, James F. Wilson, Veronique Vitart & J. Kenneth Baillie

3. Roslin Institute, University of Edinburgh, Edinburgh, UK

Erola Pairo-Castineira, Marie Zechner, Fiona Griffiths, Wilna Oosthuyzen, Jonathan Millar, Albert Tenesa, Sara Clohisey, Johnny Millar, Emma Aitkin, Ruth Armstrong, J. Kenneth Baillie, Ceilia Boz, Adam Brown, Primmy Chikowore, Judy Coyle, Louise Cullum, Nicky Day, Esther Duncan, Paul Finernan, Max Head Fourman, James Furniss, Bernadette Gallagher, Ailsa Golightly, Fiona Griffiths, Debbie Hamilton, Ross Hendry, Naomi Kearns, Dawn Law, Rachel Law, Sarah Law, Rebecca Lidstone-Scott, Hanning Mal, Ellie McMaster, Jen Meikle, Hellen Mybaya, Miranda Odam, Wilna Oosthuyzen, Nick Parkinson, Trevor Paterson, Andrew Stenhouse, Maaike Swets, Helen Szoor-McElhinney, Filip Taneski, Tony Wackett, Mairi Ward, Jane Weaver, Marie Zechner, Andrew Law, Andy Law, Sara Clohisey Hendry & J. Kenneth Baillie

4. Pain Service, NHS Tayside, Ninewells Hospital and Medical School, Dundee, UK

Andrew D. Bretherick

5. School of Life Sciences, Westlake University, Hangzhou, China

Ting Qi & Jian Yang

6. Westlake Laboratory of Life Sciences and Biomedicine, Hangzhou, China

Ting Qi & Jian Yang

7. Institute for Molecular Bioscience, The University of Queensland, Brisbane, Queensland, Australia

Yang Wu

8. Wellcome Centre for Human Genetics, University of Oxford, Oxford, UK

Isar Nassiri, Tomas Malinauskas, Julian Knight, David Stuart & Benjamin Fairfax

9. Faculty of Biological Sciences, University of Leeds, Leeds, UK

Glenn A. McConkey

10. Genomics England, London, UK

Athanasios Kousathanas

11. Intensive Care Unit, Royal Infirmary of Edinburgh, Edinburgh, UK

Jonathan Millar, Sean Keating, Manu Shankar-Hari, Timothy Walsh, J. Kenneth Baillie, Annemarie B. Docherty, Seán Keating & J. Kenneth Baillie

12. National Heart and Lung Institute, Imperial College London, London, UK

Ryan Thwaites, Peter J. M. Openshaw, Jake Dunning & Ryan S. Thwaites

13. Edinburgh Clinical Research Facility, Western General Hospital, University of Edinburgh, Edinburgh, UK

Kirstie Morrice, Richard Clark, Audrey Coutts, Lorna Donnelly, Tammy Gilchrist, Katarzyna Hafezi, Christen Lauder, Louise Macgillivray, Alan Maclean, Sarah McCafferty, Nicola Wrobel, Angie Fawkes & Lee Murphy

14. Department of Critical Care Medicine, Queen’s University and Kingston Health Sciences Centre, Kingston, Ontario, Canada

David Maslove

15. Clinical Research Centre at St Vincent’s University Hospital, University College Dublin, Dublin, Ireland

Alistair Nichol

16. NIHR Health Protection Research Unit for Emerging and Zoonotic Infections, Institute of Infection, Veterinary and Ecological Sciences University of Liverpool, Liverpool, UK

Malcolm G. Semple, Shona C. Moore, Lance Turtle, Tom Solomon, Lance C. W. Turtle & Hayley Hardwick

17. Respiratory Medicine, Alder Hey Children’s Hospital, Institute in The Park, University of Liverpool, Alder Hey Children’s Hospital, Liverpool, UK

Malcolm G. Semple

18. Centre for Inflammation Research, The Queen’s Medical Research Institute, University of Edinburgh, Edinburgh, UK

Manu Shankar-Hari, Manu Shankar-Hari, Debby Bogaert & Clark D. Russell

19. Department of Medicine, University of Cambridge, Cambridge, UK

Charlotte Summers

20. William Harvey Research Institute Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK

Charles Hinds

21. Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK

Peter Horby & Peter W. Horby

22. Department of Anaesthesia and Intensive Care, The Chinese University of Hong Kong, Prince of Wales Hospital, Hong Kong, China

Lowell Ling

23. Wellcome-Wolfson Institute for Experimental Medicine, Queen's University Belfast, Belfast, UK

Danny McAuley

24. Department of Intensive Care Medicine, Royal Victoria Hospital, Belfast, UK

Danny McAuley

25. UCL Centre for Human Health and Performance, London, UK

Hugh Montgomery

26. Imperial College Healthcare NHS Trust, London, UK

Peter J. M. Openshaw

27. Royal Hospital for Children, Glasgow, UK

Colin Begg, Fiona Bowman, Barry Milligan & Liane McPherson

28. Centre for Global Health Research, Usher Institute of Population Health Sciences and Informatics, Edinburgh, UK

Albert Tenesa & James F. Wilson

29. Genomics Division, Instituto Tecnológico y de Energías Renovables, Santa Cruz de Tenerife, Spain

Carlos Flores

30. Research Unit, Hospital Universitario N.S. de Candelaria, Santa Cruz de Tenerife, Spain

Carlos Flores, Almudena Corrales & Beatriz Guillen-Guio

31. Centre for Biomedical Network Research on Respiratory Diseases (CIBERES), Instituto de Salud Carlos III, Madrid, Spain

Carlos Flores, Almudena Corrales, Ignacio Mahillo, Patricia Muñoz García, Germán Peces-Barba & Jordi Solé-Violán

32. Department of Clinical Sciences, University Fernando Pessoa Canarias, Las Palmas de Gran Canaria, Spain

Carlos Flores & Carlos Rodriguez-Gallego

33. IDIVAL, Santander, Spain

José A. Riancho, Ana Arnaiz, María Carmen Fariñas, Ramón Fernández, Marta Fernández-Sampedro, Carmen García-Ibarbia, Silvia Martínez, J. Gonzalo Ocejo-Vinyals, Fernando Ortiz-Flores, Juan J. Ruiz-Cubillan & Miriam Vieitez-Santiago

34. Universidad de Cantabria, Santander, Spain

José A. Riancho, Ana Arnaiz, María Carmen Fariñas, Marta Fernández-Sampedro & Carmen García-Ibarbia

35. Hospital U M Valdecilla, Santander, Spain

José A. Riancho, Ana Arnaiz, María Carmen Fariñas, Marta Fernández-Sampedro, Carmen García-Ibarbia, Silvia Martínez, J. Gonzalo Ocejo-Vinyals, Fernando Ortiz-Flores, Juan J. Ruiz-Cubillan & Miriam Vieitez-Santiago

36. Tecnologico de Monterrey, Escuela de Medicina y Ciencias de la Salud and Hospital San Jose TecSalud, Monterrey, Mexico

Augusto Rojas-Martinez, Servando Cardona-Huerta, Oscar Cienfuegos-Jimenez, Jose L. Cortes-Sanchez, Emiliano Garza-Frias, Michel F. Martinez-Resendez & Rocio Ortiz-Lopez

37. Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

Pablo Lapunzina, Berta Almoguera, Carmen Ayuso, Carlos Casasnovas, Luis Castano, Marta Corton, Raquel Cruz, Silvia Diz-de Almeida, Lidia Fernandez-Caballero, Ruth Fernández-Sánchez, Inés García, Carlos Garcia-Cerrada, Rosario Lopez-Rodriguez, Eduardo López Granados, Juan José Martínez, Andrea Martínez-Ramas, Laura Marzal, Pablo Minguez, Julian Nevado, Lorena Ondo, Francesc Pla-Junca, Laura Planas-Serra, Aurora Pujol, Montserrat Ruiz, Agatha Schlüter, Marta Sevilla Porras, Cristina Silván Fuentes, Jair Antonio Tenorio Castaño, Cristina Villaverde, Miguel López de Heredia, Ingrid Mendes, Rocío Moreno, Esther Sande, Pablo Lapunzina, Angel Carracedo, Raquel Cruz & Angel Carracedo

38. Instituto de Genética Médica y Molecular (INGEMM), Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

Pablo Lapunzina, Natalia Gallego, Julian Nevado, Marta Sevilla Porras, Jair Antonio Tenorio Castaño & Pablo Lapunzina

39. ERN-ITHACA-European Reference Network, Paris, France

Pablo Lapunzina

40. Genomic Research Department, King Fahad Medical City, Riyadh, Saudi Arabia

Malak Abedalthagafi

41. Department of Pathology & Laboratory Medicine, Emory University Hospital, Atlanta, GA, USA

Malak Abedalthagafi

42. Department of Clinical Analysis and Toxicology, Federal University of Rio Grande do Norte, Natal, Brazil

Andre D. Luchessi

43. Department of Anthropology, University of Toronto at Mississauga, Mississauga, Ontario, Canada

Andre D. Luchessi & Esteban J. Parra

44. Centro Singular de Investigación en Medicina Molecular y Enfermedades Crónicas (CIMUS), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

Raquel Cruz, Silvia Diz-de Almeida, Esther Sande, Angel Carracedo, Raquel Cruz & Angel Carracedo

45. Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain

Raquel Cruz, Manuela Gago-Domínguez, Emilio Rodríguez-Ruiz, Esther Sande, Angel Carracedo & Angel Carracedo

46. Fundación Pública Galega de Medicina Xenómica, Sistema Galego de Saúde (SERGAS) Santiago de Compostela, Santiago de Compostela, Spain

Manuela Gago-Domínguez, Angel Carracedo & Angel Carracedo

47. Intensive Care National Audit & Research Centre, London, UK

Kathy Rowan

48. Heart Institute, University of Sao Paulo, Butanta, Brazil

Alexandre C. Pereira

49. NIHR Clinical Research Network (CRN), North West London Core Team, Hammersmith Hospital, London, UK

Latha Aravindan, Sukamal Das, Sheena Murphy & Mihaela Das

50. Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK

Heather Biggs, Anita Furlong, Petra Tucker & Katherine Schon

51. Biostatistics Group, State Key Laboratory of Biocontrol, School of Life Sciences, Sun Yat-sen University, Guangzhou, China

Chenqing Zheng & Jiantao Chen

52. Department of Infectious Diseases, Leiden University Medical Center, Leiden, The Netherlands

Maaike Swets

53. Guys and St Thomas' Hospital, London, UK

Jacqueline Pan, Neus Grau, Tim Owen Jones, Rosario Lim, Martina Marotti, Christopher Whitton, Aneta Bociek, Sara Campos, Gill Arbane, Manu Shankar-Hari, Marlies Ostermann, Mina Cha, Fabiola DAmato, Eirini Kosifidou, Shelley Lorah & Kyma Morera

54. James Cook University Hospital, Middlesbrough, UK

Laura Brady, Keith Hugill, Jeremy Henning, Stephen Bonner, Evie Headlam, Jessica Jones, Abigail List, Joanne Morley, Amy Welford, Bobette Kamangu, Anitha Ratnakumar & Abiola Shoremekun

55. Barts Health NHS Trust, London, UK

Zoe Alldis, Raine Astin-Chamberlain, Fatima Bibi, Jack Biddle, Sarah Blow, Matthew Bolton, Catherine Borra, Ruth Bowles, Maudrian Burton, Yasmin Choudhury, Amber Cox, Amy Easthope, Patrizia Ebano, Stavros Fotiadis, Jana Gurasashvili, Rosslyn Halls, Pippa Hartridge, Delordson Kallon, Jamila Kassam, Ivone Lancoma-Malcolm, Maninderpal Matharu, Peter May, Oliver Mitchelmore, Tabitha Newman, Mital Patel, Jane Pheby, Irene Pinzuti, Zoe Prime, Oleksandra Prysyazhna, Julian Shiel, Melanie Taylor, Carey Tierney, Olivier Zongo, Suzanne Wood, Anne Zak & David Collier

56. Royal Stoke University Hospital, Stoke-on-Trent, UK

Manuela Mundy, Christopher Thompson, Lisa Pritchard, Minnie Gellamucho, David Cartlidge, Nageswar Bandla, Lucy Bailey, Michelle Davies, Jane Delaney & Leanne Scott

57. North Middlesex University Hospital NHS Trust, London, UK

Marwa Abdelrazik, Frater Alasdair, David Carter, Munzir Elhassan, Arunkumar Ganesan, Samuel Jenkins, Zoe Lamond, Dharam Purohit, Kumar Rohit, Malik Saleem, Alanna Wall, Kugan Xavier, Dhanalaksmi Bakthavatsalam, Kirolos Gehad, Pakeerathan Gnanapragasam, Kapil Jain, Swati Jain, Abdul Malik, Naveen Pappachan, Jeronimo Moreno-Cuesta, Anne Haldeos, Rachel Vincent & Maryjane Oziegb

58. King’s College Hospital, London, UK

Anna Cavazza, Maeve Cockrell, Eleanor Corcoran, Maria Depante, Clare Finney, Ellen Jerome, Abigail Knighton, Monalisa Nayak, Evita Pappa, Rohit Saha, Sian Saha, Andrew Dodd, Kevin O’Reilly, Mark McPhail, Emma Clarey, Harriet Noble & John Smith

59. Charing Cross Hospital, St Mary’s Hospital and Hammersmith Hospital, London, UK

Phoebe Coghlan, Stephen Brett, Anthony Gordon, Maie Templeton, David Antcliffe, Dorota Banach, Sarah Darnell, Ziortza Fernandez, Eleanor Jepson, Amal Mohammed, Roceld Rojo, Sonia Sousa Arias, Anita Tamang Gurung & Jenny Wong

60. The Royal Liverpool University Hospital, Liverpool, UK

Jaime Fernandez-Roman, David O. Hamilton, Emily Johnson, Brian Johnston, Maria Lopez Martinez, Suleman Mulla, Alicia A. C. Waite, Karen Williams, Victoria Waugh, Ingeborg Welters, Jessica Emblem, Maria Norris & David Shaw

61. John Radcliffe Hospital, Oxford, UK

Archana Bashyal, Sally Beer, Paula Hutton, Stuart McKechnie, Neil Davidson, Soya Mathew, Grace Readion, Jung Ryu & Jean Wilson

62. Addenbrooke's Hospital, Cambridge, UK

Shruti Agrawal, Kay Elston, Megan Jones, Eoghan Meaney, Petra Polgarova, Muhammad Elbehery, Charlotte Summers, Esther Daubney, Anthony Ng, Jocelyn Marshall, Nazima Pathan, Katerina Stroud & Deborah White

63. Nottingham University Hospital, Nottingham, UK

Angela Andrew, Saima Ashraf, Amy Clark, Martin Dent, Margaret Langley, Cecilia Peters, Lucy Ryan, Julia Sampson, Shuying Wei, Alice Baddeley, Megan Meredith, Lucy Morris, Alexandra Gibbons & Lisa McLoughlin

64. St George's Hospital, London, UK

Carlos Castro Delgado, Victoria Clark, Deborah Dawson, Lijun Ding, Georgia Durrant, Obiageri Ezeobu, Abiola Harrison, William James Hurt, Rebecca Kanu, Ashley Kinch, Susannah Leaver, Ana Lisboa, Jisha Mathew, Kamal Patel, Romina Pepermans Saluzzio, John Rawlins, Tinashe Samakomva, Nirav Shah, Christine Sicat, Joana

Texeira, Joana Gomes De Queiroz, Edna Fernandes Da Gloria, Elena Maccacari, Nikki Yun, Soumendu Manna, Sarah Farnell-Ward, Maria Maizcordoba, Maria Thanasi & Hawakin Haji Ali

65. BHRUT (Barking Havering)—Queens Hospital and King George Hospital, Ilford, UK

Janice Hastings, Lina Grauslyte, Musarat Hussain, Bobby Ruge, Sam King, Tatiana Pogreban, Lace Rosaroso, Helen Smith, Mandeep-Kaur Phull, Nikkita Adams, George Franke, Aparna George, Erika Salciute, Joanna Wong, Karen Dunne, Luke Flower, Emma Sharland & Sukhmani Sra

66. Royal Infirmary of Edinburgh, Edinburgh, UK

Gillian Andrew, Marie Callaghan, Lucy Barclay, Lucy Marshall, Kenneth Baillie, Maria Amamio, Sophie Birch, Kate Briton, Sarah Clark, Katerine Doverman, Dave Hope, Corrienne Mcculloch, Scott Simpson & Jo Singleton

67. Kingston Hospital, London, UK

Rita Fernandez, Meryem Allen, David Baptista, Rebecca Crowe, Jonathan Fox, Jacyntha Khera, Adam Loveridge, India McKenley, Eriko Morino, Andres Naranjo, Denise O'Connor, Richard Simms, Kathryn Sollesta, Andrew Swain, Harish Venkatesh, Rosie Herdman-Grant & Anna Joseph

68. Queen Alexandra Hospital, Portsmouth, UK

Angela Nown, Steve Rose, David Pogson, Helen Boxall, Lutece Brimfield, Helen Claridge, Zoe Daly, Shenu George & Andrew Gribbin

69. Royal Gwent Hospital, Newport, UK

Yusuf Cheema, Sean Cutler, Owen Richards, Anna Roynon-Reed, Shiney Cherian, Anne Emma Heron, Gemma Williams, Tamas

Szakmany, Abby Waters, Kim Collins, Jill Dunhill, Ffion Jones, Rebecca Morris, Lucy Ship & Amy Cardwell

70. Royal Blackburn Teaching Hospital, Blackburn, UK

Syamlan Ali, Ravi Bhatterjee, Rachel Bolton, Srikanth Chukkambotla, Dabheoc Coleman, Jack Dalziel, Joseph Dykes, Christopher Fine, Bethan Gay, Wendy Goddard, Drew Goodchild, Rhiannan Harling, Muhammad Hijazi, Sarah Keith, Meherunnisa Khan, Roseanna Matt, Janet Ryan-Smith, Samuel Saad, Philippa Springle, Jacqueline Thomas, Nick Truman, Aayesha Kazi, Matthew Smith, Heather Collier, Chloe Davison, Stephen Duberley, Jeanette Hargreaves, Janice Hartley, Tahera Patel & Ellen Smith

71. Stepping Hill Hospital, Stockport, UK

Alissa Kent, Emma Goodwin, Ahmed Zaki, Clare Tibke, Susan Hopkins, Hywel Gerrard, Matthew Jackson, Sara Bennett, Liane Marsh & Rebecca Mills

72. Northumbria Healthcare NHS Foundation Trust, North Shields, UK

Jessica Bell, Helen Campbell, Angela Dawson, Steve Dodds, Stacey Duffy, Lisa Gallagher, Gemma McCafferty, Stacey Short, Tracy Smith, Kirsty Thomas, Claire Walker, Jessica Reynolds, Bryan Yates, Hayley McKie, Maria Panteli, Maria Thompson & Gail Waddell

73. Countess of Chester Hospital, Chester, UK

Sarah De Beger, Azmerelda Abraheem, Charlie Dunmore, Rumanah Girach, Rhianna Jones, Emily London, Imrun Nagra, Farah Nasir, Hannah Sainsbury, Clare Smedley, Stephen Brearey, Caroline Burchett, Kathryn Cawley, Maria Faulkner, Helen Jeffrey, Peter Bamford, Firdaus Shaikh, Lauren Slack & Angela Davies

74. Pinderfields General Hospital, Wakefield, UK

Hollie Brooke, Jose Cebrian Suarez, Ruth Charlesworth, Karen Hansson, John Norris, Alice Poole, Rajdeep Sandhu, Elizabeth Smithson, Muthu Thirumaran, Veronica Wagstaff, Sarah Buckley, Brendan Sloan, Alastair Rose, Amy Major & Alexandra Metcalfe

75. Ninewells Hospital, Dundee, UK

Christine Almaden-Boyle, Pauline Austin, Susan Chapman, Alexandre Eros, Louise Cabrelli, Stephen Cole, Clare Whyte & Matt Casey

76. Croydon University Hospital, Croydon, UK

Vasileios Bafitis, George Tsinaslanidis, Cassandra George, Reena Khade, Christopher Black & Sundar Raj Ashok

77. Morriston Hospital, Swansea, UK

Sean Farley, Elaine Brinkworth, Rachel Harford, Carl Murphy, Marie Williams, Luke Newey, Hannah Toghill, Sophie Lewis, Tabitha Rees, Ceri Battle, Mark Baker, Jenny Travers & Karen Chesters

78. Queen Elizabeth University Hospital, Glasgow, UK

Nicola Baxter, Andrew Arnott, Gordan McCreath, Christopher McParland, Laura Rooney, Malcolm Sim, Steven Henderson, Lynn Abel, Carol Dalton, Sophie Kennedy-Hay, Lynn O'Donohoe, Megan O'Hare, Izabela Orlikowska & Natasha Parker

79. Broomfield Hospital, Chelmsford, UK

Fiona McNeela, Amanda Lyle, Alistair Hughes, Jayachandran Radhakrishnan & Sian Gibson

80. Heartlands Hospital, Birmingham, UK

Hollie Bancroft, Mary Bellamy, Jacqueline Daglish, Salma Kadiri, Faye Moore, Joanne Rhodes, Mirriam Sangombe, Zhane Peterkin, James Scriven & Margaret Carmody

81. Royal Sussex County Hospital, Brighton, UK

Juliet Cottle, Emily Peasgood, Laura Ortiz-Ruiz de Gordoa, Claire Phillips & Denise Skinner

82. York Hospital, York, UK

Zoe Cinquina, Kate Howard, Rosie Joy, Samantha Roche, Isobel Birkinshaw, Joseph Carter, Jo Ingham, Nicola Marshall, Harriet Pearson & Zoe Scott

83. Queen Elizabeth Hospital, Birmingham, UK

Jo Dasgin, Jaspret Gill, Annette Nilsson, Amy Bamford, Diana Hull, James Scriven, Nafeesah Ahmadhaider, Michelle Bates & Christopher McGhee

84. Royal Glamorgan Hospital, Pontyclun, UK

Hannah Ellis, Gwenllian Sera Howe, Jayaprakash Singh, Natalie Stroud, Lisa Roche, Ceri Lynch, Bethan Deacon, Carla Pothecary, Justyna Smeaton & Kevin Agravante

85. Barnet Hospital, London, UK

Vinodh Krishnamurthy, Cynthia Diaba, Lincy John, Lai Lim & Rajeev Jha

86. Wythenshawe Hospital, Manchester, UK

Jasmine Egan, Timothy Felton, Susannah Glasgow, Grace Padden, Ozerah Choudhr, Joanne Bradley-Potts, Stuart Moss, Saejohn Lingeswaran, Peter Alexander, Craig Brandwood, Sofia Fiouni, Luke Ward, Schvearn Allen, Jane Shaw & Christopher Smith

87. Medway Maritime Hospital, Gillingham, UK

Oluronke Adanini, Rebecca Collins, Maines Msiska, Linda Ofori, Nikhil Bhatia & Hayley Dolan

88. Royal Berkshire NHS Foundation Trust, Reading, UK

Mark Brunton, Jess Caterson, Holly Coles, Liza Keating, Emma Tilney, Nicola Jacques, Matthew Frise, Jennifer Armistead, Shauna Bartley, Parminder Bhuie, Sabi Rai & Gabriela Tomkova

89. Whiston Hospital, Prescot, UK

Sandra Greer, Karen Shuker & Ascanio Tridente

90. The Royal Oldham Hospital, Manchester, UK

Emma Dobson, Jodie Hunt, Redmond Tully, Joy Dearden, Andrew Drummond, Prakash Kamath, Emily Bullock, Michelle Mulcahy, Shelia Munt, Grainne O'Connor, Jennifer Philbin, Chloe Rishton, Chloe Scott & Sarah Winnard

91. Chesterfield Royal Hospital Foundation Trust, Chesterfield, UK

Nurkamalia Hasni, Rachel Gascoyne, Joanne Hawes, Kelly Pritchard, Lesley Stevenson, Amanda Whileman, Sarah Beavis, Lauren Bishop, Cindy Cart, Katie Dale, Mary Kelly-Baxter, Adam Mendelski, Emma Moakes, Rheanna Smith, Jan Woodward & Stephanie Wright

92. Aberdeen Royal Infirmary, Aberdeen, UK

Angela Allan, Adriana Botello, Jade Liew, Jasmine Medhora, Erin Trumper, Felicity Savage, Teresa Scott, Marc Place & Callum Kaye

93. Royal Devon and Exeter Hospital, Exeter, UK

Sarah Benyon, Suzie Marriott, Linda Park, Helen Quinn, Daisy Skyes, Lily Zitter, Kizzy Baines, Elizabeth Gordon, Samantha Keenan & Andrew Pitt

94. Glasgow Royal Infirmary, Glasgow, UK

Katharine Duffy, Jane Ireland, Gary Semple, Lynne Turner, Susanne Cathcart, Dominic Rimmer, Alex Puxty, Kathryn Puxty, Andrew Hurst, Jennifer Miller, Susan Speirs & Lauren Walker

95. Blackpool Victoria Hospital, Blackpool, UK

Zena Bradshaw, Joanna Brown, Sarah Melling, Stephen Preston, Nicola Slawson, Scott Warden, Alanna Beasley, Emma Stoddard, Leonie Benham, Jason Cupitt, Melanie Caswell, Lisa Elawamy & Ashleigh Wignall

96. Southampton General Hospital, Southampton, UK

Belinda Roberts, Hannah Golding, Samantha Leggett, Michelle Male, Martyna Marani, Kirsty Prager, Toran Williams, Kim Golder, Oliver Jones, Rebecca Cusack, Clare Bolger, Rachel Burnish, Michael Carter, Susan Jackson, Karen Salmon & Jonathan Biss

97. Ashford and St Peter's Hospital, Chertsey, UK

Maia Aquino, Maria Croft, Victoria Frost, Ian White & Keshnie Govender

98. Derriford Hospital, Plymouth, UK

Natasha Webb, Liana Stapleton, Colin Wells, Nikitas Nikitas, Ana Sanchez-Rodriguez, Kayleigh Spencer & Bethan Stowe

99. East Surrey Hospital, Redhill, UK

Yvonne Izzard, Michelle Poole, Sonja Monnery, Sallyanne Trotman, Valerie Beech, Edward Combes & Teishel Joefield

100. Poole Hospital, Poole, UK

Patrick Covernton, Sarah Savage, Elizabeth Woodward, Julie Camsooksai, Henrik Reschreiter, Charlotte Barclay, Yasmin

DeAth, Judith Dube, Charlotte Humphrey, Sarah Jenkins, Emma Langridge, Rebecca Milne, Beverley Wadams & Megan Woolcock

101. Royal Alexandra Hospital, Paisley, UK

Michael Brett, Brian Digby, Lisa Gemmell, James Hornsby, Patrick MacGoey, Pauline O'Neil, Richard Price, Radha Sundaram, Lynn Abel, Natalie Rodden, Nicola Thomson, Kevin Rooney, Susan Currie, Natasha Parker, Lauren Walker & Philip Henderson

102. St James's University Hospital and Leeds General Infirmary, Leeds, UK

Bethan Ogg, Simon Whiteley, Liz Wilby, Kate Long, Shailamma Matthew, Sheila Salada, Susan Trott, Sarah Watts, Zoe Friar & Abigail Speight

103. Bedford Hospital, Bedford, UK

Victoria Bastion, Humza Chandna, Brice Djeugam, Muhammad Haseeb, Harriet Kent, Gamu Lubimbi, Sophie Murdoch, Alastair Thomas, Beena David, Rachel Lorusso, Ana Vochin, Melchizedek Penacerrada & Retno Wulandari

104. Southport and Formby District General Hospital, Ormskirk, UK

Charlotte Heath, Srinivas Jakkula, Anna Morris, Ashar Ahmed, Arvind Nune, Claire Buttriss & Emma Whitaker

105. The Tunbridge Wells Hospital and Maidstone Hospital, Kent, UK

Miriam Davey, David Golden, Amy Acklery, Fabio Fernandes, Bec Seaman & Victoria Earl

106. Queen Elizabeth Hospital, Woolwich, London, UK

Amy Collins, Waqas Khaliq, Rachel Adam & Estefania Treus

107. North Manchester General Hospital, Manchester, UK

Sarah Holland, Jordan Alfonso, Bethan Blackledge, Michelle Bruce, Laura Jayne Durrans, Ayaa Eltayeb, Jade Harris, Samuel Hey, Martin Hruska, Thomas Lamb, Joanne Rothwell, Adele Fitzgerald, Gabriella Lindergard, Helen T-Michael, Tracey Duncan, Sharon Baxter-Dore, Lisa Cooper, Claire Fox, Jacinta Guerin, Tracey Hodgkiss & Karen Connolly

108. Royal Victoria Infirmary, Newcastle Upon Tyne, UK

Paul McAlinden, Victoria Bridgett, Maggie Fearby, A. Gulati, Helen Hanson, Sinead Kelly, Louise McCormack, Rachel Nixon, Philip Robinson, Victoria Slater, Elaine Stephenson, Andrea Webster, K. Webster, Carole Hays, Anne Hudson, Bijal Patel, Ian Clement, John Davis, Sarah Francis & Douglas Jerry

109. Hull Royal Infirmary, Hull, UK

Caroline Abernathy, Louise Foster, Andrew Gratrix, Llucia Cabral-Ortega, Matthew Hines, Victoria Martinson, Elizabeth Stones & Karen Winter

110. Manchester Royal Infirmary, Manchester, UK

Esther Barrow, Katharine Wylie, Deborah Baines, Katie Birchall, Laurel Kolakaluri, Richard Clark, Anila Sukumaran, Craig Brandwood, Melanie Barker, Deborah Paripoorani, Lara Smith & Charlotte Taylor

111. Royal Derby Hospital, Derby, UK

Charlotte Downes, Melanie Hayman, Katie Riches, Priya Daniel, Deepak Subramanian, Kathleen Holding, Mary Hilton, Carly McDonald & Georgina Richardson

112. Aintree University Hospital, Liverpool, UK

Georgia Halladay, Peter Harding, Amie Reddy, Ian Turner-Bone, Laura Wilding, Robert Parker, Michaela Lloyd, Leanne Smith & Charlie

Kelly

113. Fairfield General Hospital, Bury, UK

Maria Lazo, Alan Neal, Olivia Walton, Julie Melville, Jay Naisbitt, Emily Bullock & Rosane Joseph

114. Norfolk and Norwich University Hospital (NNUH), Norwich, UK

Sara Callam, Lisa Hudig, Jocelyn Keshet-Price, Katie Stammers, Karen Convery, Georgina Randell & Deirdre Fottrell-Gould

115. Milton Keynes University Hospital, Milton Keynes, UK

Esther Mwaura, Sara-Beth Sutherland, Richard Stewart, Louise Mew & Lynn Wren

116. Good Hope Hospital, Birmingham, UK

Laura Thrasyvoulou, Heather Willis, James Scriven, Bridget Hopkins, Daniel Lenton & Abigail Roberts

117. Queen Elizabeth Hospital Gateshead, Gateshead, UK

Maria Bokhari, Rachael Lucas, Wendy McCormick, Jenny Ritzema, Vanessa Linnett, Amanda Sanderson & Helen Wild

118. Royal Bolton Hospital, Bolton, UK

Rebecca Flanagan, Robert Hull, Kat Rhead, Emma McKenna, Gareth Hughes, Jennifer Anderson, Kelly Jones, Scott Latham & Heather Riley

119. Tameside General Hospital, Ashton Under Lyne, UK

Martina Coulding, Martyn Clark, Jacqueline McCormick, Oliver Mercer, Darsh Potla, Hafiz Rehman, Heather Savill, Victoria Turner, Edward Jude & Susan Kilroy

120. Salford Royal Hospital, Manchester, UK

Elena Apetri, Cathrine Basikolo, Bethan Blackledge, Laura Catlow, Matthew Collis, Reece Doonan, Jade Harris, Alice Harvey, Karen Knowles, Stephanie Lee, Diane Lomas, Chloe Lyons, Liam McMorrow, Angiy Michael, Jessica Pendlebury, Jane Perez, Maria Poulaka, Nicola Proudfoot, Kathryn Slevin, Vicky Thomas, Danielle Walker, Paul Dark, Bethan Charles, Danielle McLaughlan, Melanie Slaughter, Dan Horner, Kathryn Cawley & Tracy Marsden

121. Great Ormond St Hospital and UCL Great Ormond St Institute of Child Health NIHR Biomedical Research Centre, London, UK

Joyann Andrews, Emily Beech, Olugbenga Akinkugbe, Alasdair Bamford, Holly Belfield, Gareth A. L. Jones, Tara McHugh, Hamza Meghari, Samiran Ray, Ana Luisa Tomas, Lauran O’Neill, Mark Peters, Michael Bell, Sarah Benkenstein, Catherine Chisholm, Charlene Davies, Klaudia Kupiec & Caroline Payne

122. Southmead Hospital, Bristol, UK

Joanna Halls, Hayley Blakemore, Elizabeth Goff, Kati Hayes, Kerry Smith, Deanna Stephens, Ruth Worner, Borislava Borislavova, Beverley Faulkner, Matt Thomas, Ruth Cookson, Emma Gendall, Georgina Larman, Rebecca Pope & Artur Smalira

123. William Harvey Hospital, Ashford, UK

Victoria Priestley, Tracey Cosier, Gemma Millen, James Rand, Natasha Schumacher, Roxana Sandhar, Heather Weston, Neil Richardson & Lucy Cooper

124. Arrowe Park Hospital, Wirral, UK

Cathy Jones, Ya-Wen Jessica Huang, Reni Jacob, Craig Denmade & Lewis McIntyre

125. Royal Hampshire County Hospital, Winchester, UK

Dawn Trodd, Jane Martin, Geoff Watson, Emily Bevan & Caroline Wreybrown

126. Bradford Royal Infirmary, Bradford, UK

Shereen Bano, Ruth Bellwood, Michael Bentley, Matt Bromley, Lucy Gurr, Camilla Ledgard, Janet McGowan, Kate Pye, Kirsten Sellick, Amelia Stacey, Deborah Warren, Brian Wilkinson, Louise Akeroyd, Huma Shafique, James Morgan, Susan Shorter, Rachel Swinger, Emily Waters & Tom Lawton

127. Glan Clwyd Hospital, Bodelwyddan, UK

Elizabeth Allan, Kate Darlington, Ffyon Davies, Llinos Davies, Jack Easton, Sumit Kumar, Richard Lean, Callum Mackay, Richard Pugh, Xinyi Qiu, Stephanie Rees, Jeremy Scanlon, Joanne Lewis, Daniel Menzies, Annette Bolger, Gwyneth Davies, Jennifer Davies, Esther Garrod, Helen Jones, Rachel Manley & Hannah Williams

128. Royal Bournemouth Hospital, Bournemouth, UK

Jordan Frankham, Sally Pitts, Nigel White, Debbie Branney & Heather Tiller

129. Bristol Royal Infirmary, Bristol, UK

Georgia Efford, Zoe Garland, Lisa Grimmer, Bethany Gumbrill, Rebekah Johnson, Katie Sweet, Jeremy Bewley, Christina Coleman, Katie Corcoran, Eva Maria Hernandez Morano, Rachel Shiel, Denise Webster, Josephine Bonnici, Eleanor Daniel & Abbie Dell

130. University Hospital North Durham, Darlington, UK

Melanie Kent, Ami Wilkinson, Ellen Brown, Andrea Kay, Suzanne Campbell, Amanda Cowton, Mark Birt, Vicki Greenaway, Kathryn

Potts, Clare Hutton & Andrew Shepperson

131. Darlington Memorial Hospital, Darlington, UK

Melanie Kent, Ami Wilkinson, Ellen Brown, Andrea Kay, Suzanne Campbell, Amanda Cowton, Mark Birt, Vicki Greenaway, Kathryn Potts, Clare Hutton & Andrew Shepperson

132. Basildon Hospital, Basildon, UK

Miranda Forsey, Alice Nicholson, Mark Vertue, Joanne Riches, Agilan Kaliappan & Anne Nicholson

133. University College Hospital, London, UK

Niall MacCallum, Eamon Raith, Georgia Bercades, Ingrid Hass, David Brealey, Gladys Martir, Anna Reyes, Deborah Smyth & Maria Zapatamartinez

134. Whittington Hospital, London, UK

Ana Alvaro, Champa Jetha, Louise Ma, Lauren Booker, Loreta Mostoles, Anezka Pratley, Abdelhakim Altabaibeh, Chetan Parmar & Kayleigh Gilbert

135. Western General Hospital, Edinburgh, UK

Susie Ferguson, Amy Shepherd, Sheila Morris, Jo Singleton, Rosie Baruah, Maria Amamio, Sophie Birch, Kate Briton, Sarah Clark, Katherine Doverman, Lucy Marshall & Scott Simpson

136. Ipswich Hospital, Ipswich, UK

Georgina Lloyd, Stephanie Bell, Vanessa Rivers & Bally Purewal

137. Hereford County Hospital, Hereford, UK

Kate Hammerton, Susan Anderson, Janine Birch, Emma Collins & Ryan Oleary

138. Sunderland Royal Hospital, Sunderland, UK

Sarah Cornell, Jordan Jarmain, Kimberley Rogerson, Fiona Wakinshaw, Lindsey Woods, Anthony Rostron, Zeynep Elcioglu & Alistair Roy

139. Queens Hospital Burton, Burton-On-Trent, UK

Gillian Bell, Holly Dickson, Louise Wilcox, Amro Katary & Katy English

140. Musgrove Park Hospital, Taunton, UK

Joanne Hutter, Corinne Pawley, Patricia Doble, Charmaine Shovelton, Marius Vaida, Rebecca Purnell & Ashly Thomas

141. The Royal Papworth Hospital, Cambridge, UK

Lenka Cagova, Adama Fofano, Helen Holcombe, Alice Michael Mitchell, Lucy Mwaura, Krithivasan P. Raman, Lucie Garnr, Sue Mepham, Kitty Paques, Alain Vuylsteke, Jennifer Mackie, Carmen Pearn & Julie Zamikula

142. University Hospital Lewisham, London, UK

Mark Birt, Estefania Treus Gude, Maggie Nyirenda, Lisa Capozzi, Rosie Reece-Anthony, Waqas Khaliq, Hazma Noor & Alfa Cresia Nilo

143. The Princess Alexandra Hospital, Harlow, UK

Michelle Grove, Amelia Daniel, Amy Easthope, Joanne Finn, Nikki White, Rajnish Saha, Bibi Badal & Karen Ixer

144. University Hospital of Wales, Cardiff, UK

Donna Duffin, Ben Player, Helen Hill, Jade Cole, Jenny Brooks, Michelle Davies, Rhys Davies, Lauren Hunt, Emma Thomas & Angharad Williams

145. West Middlesex Hospital, Isleworth, UK

Metod Oblak, Mini Thankachen, Jamie Irisari, Amrinder Sayan & Monica Popescu

146. Royal Albert Edward Infirmary, Wigan, UK

Cheryl Finch, Andrew Jamieson, Alison Quinn, Joshua Cooper, Sarah Liderth & Natalia Waddington

147. Stoke Mandeville Hospital, Aylesbury, UK

Iona Burn, Katarina Manso, Ruth Penn, Julie Tebbutt, Danielle Thornton, James Winchester, Geraldine Hambrook & Pradeep Shanmugasundaram

148. Royal Lancaster Infirmary, Lancaster, UK

Jayne Craig, Kerry Simpson, Andrew Higham & Louise Sibbett

149. Basingstoke and North Hampshire Hospital, Basingstoke, UK

Sheila Paine, Annabel Reed, Jo-Anna Conyngham, McDonald Mupudzi, Rachel Thomas, Mary Wright, Denise Griffin, Richard Partridge, Maria Alvarez Corral, Nycola Muchenje, Mildred Sitonik & Caroline Wrey Brown

150. Worthing Hospital, Worthing, UK

Aaron Butler, Linda Folkes, Heather Fox, Amy Gardner, David Helm, Gillian Hobden, Kirsten King, Jordi Margalef, Michael Margarson, Tim Martindale, Emma Meadows, Dana Raynard, Yvette Thirlwall, Yolanda Baird, Raquel Gomez, Darren Martin, Luke Hodgson, Clinton Corin, Erikka Sidall, Densie Szabo & Sharon Floyd

151. St Richard’s Hospital, Chichester, UK

Aaron Butler, Linda Folkes, Heather Fox, Amy Gardner, David Helm, Gillian Hobden, Kirsten King, Jordi Margalef, Michael

Margarson, Tim Martindale, Emma Meadows, Dana Raynard, Yvette Thirlwall, Yolanda Baird, Raquel Gomez, Darren Martin, Luke Hodgson, Clinton Corin, Erikka Sidall, Densie Szabo & Sharon Floyd

152. The Alexandra Hospital, Redditch and Worcester Royal Hospital, Worcester, UK

Hannah Davies, Karen Austin, Olivia Kelsall, Hannah Wood, Peter Anderson, Katie Archer, Andrew Burtenshaw, Sarah Clayton, Naiara Cother, Nicholas Cowley, Caroline Davis, Stephen Digby, Alison Durie, Alison Harrison, Emma Low, Michael McAlindon, Alex McCurdy, Aled Morgan, Tobias Rankin, Jessica Thrush, Helen Tranter, Charlie Vigurs & Laura Wild

153. Royal Cornwall Hospital, Truro, UK

Thomas Cornell, Kate Ralph, Sarah Bean, Karen Burt, Michael Spivey, Carol Richards & Rachel Tedstone

154. Watford General Hospital, Watford, UK

Siobhain Carmody, Xiaobei Zhao, Valerie Page, Mark Louie Guanco, Elvira Hoxha & Camilla Zorloni

155. Macclesfield District General Hospital, Macclesfield, UK

Charlotte Dean, Emma Jones, Emma Carter, Joshua Dunn, Thomas Kong, Mervin Mahenthran, Chris Marsh, Maureen Holland, Natalie Keenan, Mohamed Mahmoud, Marc Lyons, Joanne Bradley-Potts, Helen Wassall & Meghan Young

156. Royal Surrey County Hospital, Guildford, UK

Paul Bradley, Dorota Burda, Sinead Donlon, Lesley Harden, Celia Harris, Irving Mayangao, Rugia Montaser, Sheila Mtuwa, Charles Piercy, Eleanor Smith, Sarah Stone, Jerik Verula, Helen Blackman, Cheryl Marriott, Natalia Michalak, Ben Creagh-Brown, Armorel Salberg, Naomi Boyer & Veronika Pristopan

157. Rotherham General Hospital, Rotherham, UK

Victoria Maynard, Rachel Walker, Anil Hormis, Dawn Collier, Cheryl Graham, Vicky Maynard, Jake McCormick & Jake Warrington

158. Craigavon Area Hospital, Portadown, UK

Denise Cosgrove, Denise McFarland, Judith Ratcliffe & Rob Charnock

159. King's Mill Hospital, Nottingham, UK

Inez Wynter, Mandy Gill, Jill Kirk, Paul Paul, Valli Ratnam & Sarah Shelton

160. Dumfries and Galloway Royal Infirmary, Dumfries, UK

Catherine Jardine, Alasdair Hay & Dewi Williams

161. Prince Charles Hospital, Merthyr Tydfil, UK

Bethan Deacon, Latha Durga, Meg Hibbert, Gareth Kennard-Holden, Chrsitopher Woodford, Carla Pothecary, Lisa Roche, Dariusz Tetla, Kevin Agravante & Justyna Smeaton

162. Ysbyty Gwynedd, Bangor, UK

Alicia Price, Alice Thomas, Chris Thorpe, Ellen Knights & Donna Ward

163. Royal Preston Hospital, Preston, UK

Shondipon Laha, Mark Verlander & Alexandra Williams

164. The Great Western Hospital, Swindon, UK

Rachel Prout, Helen Langton, Malcolm Watters, Charlotte Hunt & Catherine Novis

165. Lincoln County Hospital, Lincoln, UK

Sarwat Arif, Amy Cunningham, Claire Hewitt, Julia Hindale, Karen Jackson-Lawrence, Sarah Shepardson, Maryanne Wills, Susie Butler, Silivia Tavares, Russell Barber, Annette Hilldrith & Kelly Hubbard

166. University Hospital of North Tees, Stockton on Tees, UK

Dawn Egginton, Michele Clark, Sarah Purvis, Simon Sinclair & Vicky Collins

167. Glangwili General Hospital, Camarthen, UK

Bethan Landeg, Craig Sell, Samantha Coetzee, Alistair Gales, Igor Otahal, Becky Icke, Meena Raj, Caroline Williams, Jill Williams & Lucy Hill

168. Southend University Hospital, Westcliff-on-Sea, UK

Abdul Kayani, Bridgett Masunda, Prisca Gondo & Nigara Atayeva

169. Lister Hospital, Stevenage, UK

Carina Cruz & Natalie Pattison

170. Diana Princess of Wales Hospital, Grimsby, UK

Caroline Burnett, Jonathan Hatton, Elaine Heeney, Maria Newton, Hassan Al-Moasseb, Teresa Behan, Jasmine Player, Rachael Stead, Atideb Mitra & Kirsty Nauyokas

171. West Suffolk Hospital, Bury St Edmunds, UK

Sally Humphreys, Helen Cockerill & Ruth Tampsett

172. Victoria Hospital, Kirkcaldy, UK

Evgeniya Postovalova, Tina Coventry, Amanda McGregor, Susan Fowler, Mike Macmahon, Patricia Cochrane & Sandra Pirie

173. Calderdale Royal Hospital, Halifax, UK

Sarah Hanley, Asifa Ali, Megan Brady, Sam Dale, Annalisa Dance, Lisa Gledhill, Jill Greig, Kathryn Hanson, Kelly Holdroyd, Marie Home, Tahira Ishaq, Diane Kelly, Lear Matapure, Deborah Melia, Samantha Mellor, Ekta Merwaha, Tonicha Nortcliffe, Lisa Shaw, Ryan Shaw, Tracy Wood, Lee-Ann Bayo, Miranda Usher, Alison Wilson, Ross Kitson, Jez Pinnell, Matthew Robinson & Kaitlin Boltwood

174. Huddersfield Royal Infirmary, Huddersfield, UK

Sarah Hanley, Asifa Ali, Megan Brady, Sam Dale, Annalisa Dance, Lisa Gledhill, Jill Greig, Kathryn Hanson, Kelly Holdroyd, Marie Home, Tahira Ishaq, Diane Kelly, Lear Matapure, Deborah Melia, Samantha Mellor, Ekta Merwaha, Tonicha Nortcliffe, Lisa Shaw, Ryan Shaw, Tracy Wood, Lee-Ann Bayo, Miranda Usher, Alison Wilson, Ross Kitson, Jez Pinnell, Matthew Robinson & Kaitlin Boltwood

175. Dorset County Hospital, Dorchester, UK

Jenny Birch, Laura Bough, Rebecca Tutton, Barbara Winter-Goodwin, Josie Goodsell, Kate Taylor, Patricia Williams, Sarah Williams, Ashleigh Cave & James Rees

176. Russell's Hall Hospital, Dudley, UK

Janet Imeson-Wood, Jacqueline Smith, Vishal Amin, Komala Karthik, Rizwana Kausar, Elena Anastasescu, Karen Reid, Vikram Anumakonda & Ella Stoddart

177. Royal United Hospital, Bath, UK

Carrie Demetriou, Charlotte Eckbad, Lucy Howie, Sarah Mitchard, Lidia Ramos, Katie White, Sarah Hierons, Fiona Kelly, Alfredo Serrano-Ruiz & Gabrielle Evans

178. St Mary’s Hospital, Newport, UK

Liz Nicol, Joy Wilkins, Kim Hulacka, Gabor Debreceni, Alison Brown & Vikki Crickmore

179. George Eliot Hospital NHS Trust, Nuneaton, UK

Kay Hill & Thogulava Kannan

180. Yeovil Hospital, Yeovil, UK

Zenaida Dagutao, Kate Beesley, Alison Lewis, Jess Perry, Sherly Antony, Sarah Board, Clare Buckley, Lucy Pippard, Alfonso Tanate, Diane Wood, Agnieska Kubisz-Pudelko & Ayman Gouda

181. Forth Valley Royal Hospital, Falkirk, UK

Fiona Auld, Joanne Donnachie, Euan Murdoch, Lynn Prentice, Nikole Runciman, Dhaneesha Senaratne, Abigail Short, Laura Sweeney, Lesley Symon, Anne Todd, Patricia Turner, Erin McCann, Dario Salutous, Ian Edmond & Lesley Whitelaw

182. Frimley Park Hospital, Camberley, UK

Harish Venkatesh, Yvonne Bland, Istvan Kajtor, Lisa Kavanagh, Karen Singler & George Linfield-Brown

183. Chelsea & Westminster NHS Foundation Trust, London, UK

Luke Stephen Prockter Moore, Marcela Vizcaychipi, Laura Martins, Luke Moore, Rhian Bull & Jaime Carungcong

184. Queen Elizabeth the Queen Mother Hospital, Margate, UK

Louise Allen, Eva Beranova, Alicia Knight, Carly Price, Sorrell Tilbey, Sharon Turney, Tracy Hazelton, Gabriella Tutt, Mansi Arora, Salah Turki, Emily Sinfield, Joanne Deery & Hazel Ramos

185. Royal Brompton Hospital, London, UK

Daniele Cristiano, Natalie Dormand, Zohreh Farzad, Mahitha Gummadi, Sara Salmi, Geraldine Sloane, Mathew Varghese, Vicky Thwaites, Brijesh Patel, Liyanage Kamal & Anelise Catelan Zborowski

186. Darent Valley Hospital, Dartford, UK

Ryan Coe, Madeleine Anderson, Jane Beadle, Charlotte Coates, Katy Collins, Maria Crowley, Laura Johnson, Laura King, Remi Paramsothy, Janet Sargeant, Pedro Silva, Carmel Stuart, June Taylor, David Tyl, Phillipa Wakefield, Charlotte Kamundi, Olumide Olufuwa, Zakaulla Belagodu, Anca Gherman & Naomi Oakley

187. University Hospital Crosshouse, Kilmarnock, UK

John Allan, Tim Geary, Alistair Meikle, Peter O'Brien, Stephen Wood, Andrew Clark & Gordon Houston

188. University Hospital Wishaw, Wishaw, UK

Karen Black, Michelle Clarkson, Stuart D'Sylva, Alan Morrison, Kathryn Norman, Margaret Taylor, Suzanne Clements, Catriona Cohrane, Nora Gonzalez, Dominic Strachan, Claire Beith & Kirsten Moar

189. University College Dublin, St Vincent's University Hospital, Dublin, Ireland

Lorna Murphy, Michelle Smythe, Alistair Nichol & Kathy Brickell

190. The Queen Elizabeth Hospital, King's Lynn, UK

Inthakab Ali Mohamed Ali, Karen Beaumont, Mohamed Elsaadany, Kay Fernandes, Sameena Mohamed Ally, Harini Rangarajan, Varun Sarathy, Sivarupan Selvanayagam, Dave Vedage, Matthew White, Zoe Coton, Aricsa Joshy, Mark Blunt & Hollie Curgenven

191. Walsall Manor Hospital, Walsall, UK

Liam Botfield, Catherine Dexter, Aditya Kuravi, Joanne Butler, Robert Chadwick, Poonam Ranga, Lisa Richardson, Emma Virgilio, Maddiha Anwer, Atul Garg, Donna Botfield & Xana Marriott

192. Princess Royal Hospital, Brighton, UK

Keely Stewart, Dee Mullan, Claire Phillips, Jane Gaylard, Justyna Nowak & Denise Skinner

193. Barnsley Hospital, Barnsley, UK

Sian Jones, Rikki Crawley, Abigail Crew, Mishell Cunningham, Allison Daniels, Laura Harrison, Susan Hope, Nicola Lancaster, Jamie Matthews, Gemma Wray, Alice Nicholson, Ken Inweregbu, Sarah Cutts & Katharine Miller

194. Warrington General Hospital, Warrington, UK

Ailbhe Brady, Rebekah Chan, Shane McIvor, Helena Prady, Bijoy Mathew, Jeff Little & Tim Furniss

195. Royal Victoria Hospital, Belfast, UK

Chris Wright, Bernadette King, Christopher Wasson, Aisling O'Neill, Christine Turley, Peter McGuigan, Erin Collins, Stephanie Finn, Jackie Green, Julie McAuley, Abitha Nair, Charlotte Quinn, Suzanne Tauro, Kathryn Ward, Michael McGinlay & Kiran Reddy

196. Royal Hallamshire Hospital and Northern General Hospital, Sheffield, UK

Norfaizan Ahmad, Samantha Anderson, Joann Barker, Kris Bauchmuller, Kathryn Birchall, Sarah Bird, Kay Cawthron, Luke Chetam, Joby Cole, Ben Donne, David Foote, Amber Ford, Helena Hanratty, Kate Harrington, Lisa Hesseldon, Kay Housley, Yvonne Jackson, Claire Jarman, Faith Kibutu, Becky Lenagh, Irene Macharia, Shamiso Masuko, Leanne Milner, Helen Newell, Lorenza

Nwafor, Simon Oxspring, Patrick Phillips, Ajay Raithatha, Sarah Rowland-Jones, Jacqui Smith, Roger Thompson, Helen Trower, Sara Walker, James Watson, Matthew Wiles, Alison Lye, Jayne Willson, Gary Mills, Sansha Harris & Eleanor Hartill

197. Harefield Hospital, London, UK

Anthony Barron, Ciara Collins, Sundeep Kaul, Claire Nolan, Oliver Polgar, Claire Prendergast, Paula Rogers, Rajvinder Shokkar, Meriel Woodruff, Kanta Mahay, Vicky Thwaites, Anna Reed, Hayley Meyrick, Heather Passmore & James Farwell

198. Cumberland Infirmary, Carlisle, UK

Alison Brown, Susan O'Connell, Jane Gregory, Luigi Barberis, Rosemary Harper, Tim Smith & Diane Armstrong

199. Eastbourne District General Hospital, Eastbourne, UK

Angie Bowey, Anne Cowley, Andrew Corner, Judith Highgate, Claire Rutherfurd, Jo-Anne Taylor, Sarah Goodwin & Claire Rutherford

200. Conquest Hospital, Saint Leonards-on-Sea, UK

Angie Bowey, Anne Cowley, Andrew Corner, Judith Highgate, Claire Rutherfurd, Jo-Anne Taylor, Sarah Goodwin & Claire Rutherford

201. Salisbury District Hospital, Salisbury, UK

Beena Eapen, Fiona Trim & Phil Donnison

202. Airedale General Hospital, Keighley, UK

Lisa Armstrong, Hayley Bates, Emma Dooks, Fiona Farquhar, Amy Kitching, Chantal McParland, Sophie Packham & Brigid Hairsine

203. Leicester Royal Infirmary, Leicester, UK

Premetie Andreou, Dawn Hales, Megha Mathews, Rekha Patel, Peter Barry, Neil Flint, Jessica Hailstone, Navneet Ghuman, Bethany Leonard & Rachel Lees

204. Peterborough City Hospital, Peterborough, UK

Deborah Butcher, Katy Leng, Nicola Butterworth-Cowin & Susie O'Sullivan

205. Hinchingbrooke Hospital, Huntingdon, UK

Deborah Butcher, Katy Leng, Nicola Butterworth-Cowin & Susie O'Sullivan

206. Colchester General Hospital, Colchester, UK

Alison Ghosh & Emma Williams

207. Princess Royal Hospital, Telford and Royal Shrewsbury Hospital, Shrewsbury, UK

Colene Adams, Anita Agasou, Tracie Arden, Mandy Beekes, Amy Bowes, Pauline Boyle, Heather Button, Mandy Carnahan, Anne Carter, Danielle Childs, Jane Gaylard, Fran Hurford, Yasmin Hussain, Ayesha Javaid, James Jones, Michael Leigh, Terry Martin, Helen Millward, Nichola Motherwell, Dee Mullan, Julie Newman, Rachel Rikunenko, Jo Stickley, Julie Summers, Louise Ting, Helen Tivenan, Denise Donaldson, Nigel Capps, Emily Cale, Sanal Jose, Wendy Osbourne, Susie Pajak, Jayne Rankin & Louise Tonks

208. University Hospital Monklands, Airdrie, UK

Tracy Baird, Margaret Harkins, Jim Ruddy & Joe West

209. Wrexham Maelor Hospital, Wrexham, UK

Joseph Duffield, Lewis Mallon, Oliver Smith, Sara Smuts, Andy Campbell, Cate Davies, Sarah Davies, Rachel Hughes, Lisa

Jobes, Victoria Whitehead & Clare Watkins

210. New Cross Hospital, Wolverhampton, UK

Stella Metherell, Nichola Harris, Victoria Lake, Elizabeth Radford, Andy Smallwood, Shameer Gopal & Katherine Vassell

211. University Hospital Hairmyres, East Kilbride, UK

Dina Bell, Rosalind Boyle, Katie Douglas, Lynn Glass, Liz Lennon, Austin Rattray, Claire Beith & Emma Lee

212. Warwick Hospital, Warwick, UK

Danielle Jones, Penny Parsons, Ben Attwood, Paul Jefferson, Mohan Ranganathan, Inderjit Atwal, Bridget Campbell, Angela Day & Camilla Stagg

213. Sandwell General Hospital and City Hospital, Birmingham, UK

Emma Haynes, Cecilia Ahmed, Sarah Clamp, Julie Colley, Risna Haq, Anne Hayes, Sibet Joseph, Zahira Maqsood, Samia Hussain, Jonathan Hulme, Patience Domingos, Rita Kumar, Manjit Purewal & Becky Taylor

214. Royal Manchester Children's Hospital, Manchester, UK

Lara Bunni, Monica Latif, Claire Jennings, Shilu Jose, Rebecca Marshall, Aleksandra Metryka & Gayathri Subramanian

215. Gloucestershire Royal Hospital, Gloucester, UK

Adam Burgoyne, Susan O'Connell, Amanda Tyler, Joanne Waldron, Paula Hilltout & Jayne Evitts

216. University Hospitals Coventry & Warwickshire NHS Trust, Coventry, UK

Geraldine Ward, Pamela Bremmer, Carl Hawkins, Sophie Jackman & Michal Ogorek

217. Torbay Hospital, Torquay, UK

Kylie Ashby, Lorraine Thornton, Pauline Mercer, Matthew Halkes & Adam Revill

218. Pilgrim Hospital, Lincoln, UK

Bryony Saint, Jo Fletcher, Kimberley Netherton, Manish Chablani, Amy Kirkby, Amanda Roper & Kinga Szymiczek

219. Prince Philip Hospital, Lianelli, UK

Isobel Sutherland, Linda O’Brien, Igor Otahal, Joanne Connell, Kim Davies, Tracy Lewis, Zohra Omar & Emma Perkins

220. Princess of Wales Hospital, Llantrisant, UK

Lisa Roche, Sonia Sathe & Ellie Davies

221. Northampton General Hospital NHS Trust, Northampton, UK

Alex Lyon, Isheunesu Mapfunde, Charlotte Willis, Rachael Hitchcock, Kathryn Hall & Christopher King

222. The Christie NHS Foundation Trust, Manchester, UK

Andrew Fagan, Roonak Nazari, Lucy Worsley, Suzanne Allibone, Vidya Kasipandian, Amit Patel, Parisa Cutting, Roman Genetu, Ainhi Mac, Anthony Murphy, Sinead Ward & Fatima Butt

223. James Paget University Hospital NHS Trust, Great Yarmouth, UK

Amanda Ayers, Wendy Harrison, Katherine Mackintosh & Julie North

224. Birmingham Children’s Hospital, Birmingham, UK

Lydia Ashton, Rehana Bi, Samantha Owen, Helen Winmill & Barney Scholefield

225. Withybush General Hospital, Haverfordwest, UK

Hannah Blowing, Erin Williams, Michaela Duskova, Michelle Edwards, Alun Rees, Helen Thomas, Rachel Hughes, Igor Otahal, Jolene Brooks, Janet Phipps & Suzanne Brooks

226. Northwick Park Hospital, London, UK

Catherine Dennis, Vicki Parris, Sinduya Srikaran, Anisha Sukha, Alistair McGregor & Gerlynn Tiongson

227. North Devon District Hospital, Barnstaple, UK

Katie Adams, Benedict Andrew, Adam Brayne, Sasha Carter, Louise Findlay, Emma Fisher, Peter Jackson, Duncan Kaye, Juliet Parkin, Victoria Tuckey, Jane Hunt, Nicholas Love, Lynne van Koutrick & Ashley Hanson

228. Scunthorpe General Hospital, Scunthorpe, UK

Kathy Dent, Elizabeth Horsley, Sandra Pearson, Sue Spencer, Dorothy Hutchinson, Jasmine Player, Dorota Potoczna, Muhammad Nauman Akhtar, Lisa-Jayne Cottam, Kirsty Nauyokas & Jack Sanders

229. Royal Free Hospital, London, UK

Sara Mingo Garcia, Glykeria Pakou, Cynthia Diaba, Helder Filipe, Lincy John, Amitaa Maharajh, Mark de Neef, Daniel Martin, Christine Eastgate & Poh Choo Teoh

230. Raigmore Hospital, Inverness, UK

Fiona Barrett, Clare Bradley, Avril Donaldson, Mairi Mascarenhas, Marianne O’Hara, Laura Okeefe, Noreen Clarke, Jonathan Whiteside, Rachael Campbell, Joanna Matheson, Deborah McDonald & Donna Patience

231. West Cumberland Hospital, Whitehaven, UK

Polly Rice, Tim Smith, Melanie Clapham, Rachel Mutch, Luigi Barberis, Rosemary Harper, Hannah Craig & Una Poultney

232. Furness General Hospital, Barrow-in-Furness, UK

Karen Burns & Andrew Higham

233. Liverpool Heart and Chest Hospital, Liverpool, UK

Sophie Twiss, Janet Barton, Linsha George, Clare Harrop, Sherly Mathew & David Justin Wright

234. Scarborough General Hospital, Scarborough, UK

Rachel Harrison, Jordan Toohie, Ben Chandler, Alison Turnbull, Janine Mallinson & Kerry Elliott

235. Bronglais General Hospital, Aberystwyth, UK

Rebecca Wolf-Roberts, Helen Tench, Igor Otahal, Maria Hobrok, Ronda Loosley, Heather McGuinness & Tanya Sims

236. Alder Hey Children's Hospital, Liverpool, UK

Deborah Afolabi, Kathryn Sian Allison, Taya Anderson, Rachael Dore, Dawn Jones, Naomi Rogers, Paula Saunderson, Jennifer Whitbread, Laura O'Malley, Laura Rad & Daniel Hawcutt

237. Borders General Hospital, Melrose, UK

Jonathan Aldridge, Melanie Tolson & Sweyn Garrioch

238. Leighton Hospital, Crewe, UK

Joanne Tomlinson & Michael Grosdenier

239. Kent & Canterbury Hospital, Canterbury, UK

David Loader, Ritoo Kapoor & Gemma Hector

240. Harrogate and District NHS Foundation Trust, Harrogate, UK

Joslan Scherewode, Chunda Sri-Chandana, Lorraine Stephenson & Sarah Marsh

241. The Royal Marsden Hospital, London, UK

Arnold Dela Rosa, Shaman Jhanji, Thomas Bemand, Ryan Howle, Ravishankar Rao Baikady, Benjamin Thomas, Ethel Black & Kate Tatham

242. Ealing Hospital, Southall, UK

Sambasivarao Gurram, Ekaterina Watson, Vicki Parris, Sheena Quaid & Alistair McGregor

243. St John's Hospital Livingston, Livingston, UK

Anne Saunderson, Rachel O'Brien, Sam Moultrie, Jen Service, Clare Cheyne, Miranda Odam & Alison Wiliams

244. Wexham Park Hospital, Slough, UK

Nicky Barnes, Peter Csabi, Joana Da Rocha & Louika Glynou

245. Sheffield Children's Hospital, Sheffield, UK

Amy Huffenberger, Jade Bryant, Amy Pickard, Nicholas Roe, Arianna Bellini, Anton Mayer, Amy Burrow, Natalie Colley, Jayne Evans, Alex Howlett & Zeinab Khalifeh

246. Homerton University Hospital Foundation NHS Trust, London, UK

Jerldine Pryce, Claire Gorman, Amy Easthope, Rebecca Brady, Elizabeth Timlick, Pierre Antoine & Abhinhav Gupta

247. National Hospital for Neurology and Neurosurgery, London, UK

John Hardy, Henry Houlden, Eleanor Moncur, Arianna Tucci, Eamon Raith, Ambreen Tariq & David Brealey

248. The Royal Alexandra Children's Hospital, Brighton, UK

Emma Tagliavini, Becky Ramsay, Katy Fidler, Kevin Donnelly & Rebecca Hollis

249. Golden Jubilee National Hospital, Clydebank, UK

Jocelyn Barr, Elizabeth Boyd, Val Irvine, Ben Shelley, Julie Buckley, Charlene Hamilton & Kathryn Valdeavella

250. Hospital Universitario Mostoles, Medicina Interna, Madrid, Spain

Javier Abellan, Carlos Garcia-Cerrada, Victor Moreno Cuerda, Belén Rodríguez Maya & Jorge Sánchez Redondo

251. Universidad Francisco de Vitoria, Madrid, Spain

Javier Abellan, Carlos Garcia-Cerrada, María Carmen García Torrejón & Victor Moreno Cuerda

252. Haemostasis and Thrombosis Unit, Hospital de la Santa Creu i Snt Pau, IIB Sant Pau, Barcelona, Spain

René Acosta-Isaac & Juan Carlos Souto

253. Unit of Infectious Diseases, Hospital Universitario 12 de Octubre, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain

Jose María Aguado

254. Spanish Network for Research in Infectious Diseases (REIPI RD16/0016/0002), Instituto de Salud Carlos III, Madrid, Spain

Jose María Aguado

255. School of Medicine, Universidad Complutense, Madrid, Spain

Jose María Aguado, Álvaro Andreu-Bernabeu, Celso Arango, Covadonga M. Diaz-Caneja, Javier González-Peñas, Patricia Muñoz García & Mara Parellada

256. Centro de Investigación Biomédica en Red de Enfermedades Infecciosas (CIBERINFEC), Instituto de Salud Carlos III, Madrid, Spain

Jose María Aguado, Óscar Brochado-Kith, Amanda Fernández-Rodríguez, Juan P. Horcajada, María A. Jimenez-Sousa & Salvador Resino

257. Hospital General Santa Bárbara de Soria, Soria, Spain

Carlos Aguilar & Fernando Sevil Puras

258. Pediatric Neurology Unit, Department of Pediatrics, Navarra Health Service Hospital, Pamplona, Spain

Sergio Aguilera-Albesa

259. Navarra Health Service, NavarraBioMed Research Group, Pamplona, Spain

Sergio Aguilera-Albesa

260. Complejo Asistencial Universitario de León, León, Spain

Abdolah Ahmadi Sabbagh, Belén Ballina Martín, Silvia Fernández Ferrero, Yolanda Fernández Martínez, Marta Fuertes Núñez, Cristina Hernández Moro, Violeta Martínez Robles, Irene Padilla Conejo, Lisbeth A. Pichardo, José A. Rodriguez-Garcia, Filomeno Rondón García, Javier Sánchez Real, Julia Vidán Estévez & Lavinia Villalobos

261. Infectious Diseases Department, Hospital Universitario San Pedro, Logroño, Spain

Jorge Alba & José A. Oteo

262. Fundación Institut Guttmann, Institut Universitari de Neurorehabilitació adscrit a la UAB, Hospital de Neurorehabilitació, Barcelona, Spain

Sergiu Albu

263. Universitat Autònoma de Barcelona (UAB), Barcelona, Spain

Sergiu Albu & Juan P. Horcajada

264. Fundació Institut d'Investigació en Ciències de la Salut Germans Trias i Pujol, Barcelona, Spain

Sergiu Albu

265. Hospital General de Occidente, Guadalajara, Mexico

Karla A. M. Alcalá-Gallardo & Marco A. Cid-Lopez

266. Microbiology Unit, Hospital Universitario N.S. de Candelaria, Santa Cruz de Tenerife, Spain

Julia Alcoba-Florez

267. Servicio de Neumología, Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

Sergio Alcolea Batres, Rodolfo Alvarez-Sala Walther, Carlos Carpio Segura, Luis Gómez Carrera, Daniel Laorden, Pablo Mariscal Aguilar & María Concepción Sánchez Prados

268. Camino Universitario Adelita de Char, Mired IPS, Barranquilla, Colombia

Holmes Rafael Algarin-Lara, Yady Álvarez-Benítez, Sylena Chiquillo-Gómez & María Eugenia Quevedo Chávez

269. Facultad de Ciencias de la Salud, Universidad Simón Bolívar, Barranquilla, Colombia

Holmes Rafael Algarin-Lara, Yady Álvarez-Benítez, Andrea Barranco-Díaz, Sylena Chiquillo-Gómez, Anderson Díaz-Pérez, Lácides Fuenmayor-Hernández, Humberto Mendoza Charris, Xenia Morelos-Arnedo, Junior Moreno-Escalante, Fredy Javier Pacheco-Miranda, María Eugenia Quevedo Chávez, Marena Rodríguez-Ferrer, Andrea Romero-Coronado & Zuleima Yáñez

270. Neumología, Hospital Universitario Virgen Macarena, Seville, Spain

Virginia Almadana & Agustín Valido

271. Departamento de Medicina, Universidad de Salamanca, Salamanca, Spain

Julia Almeida & Alberto Orfao

272. Centro de Investigación del Cáncer (IBMCC), Universidad de Salamanca, CSIC, Salamanca, Spain

Julia Almeida & Alberto Orfao

273. Biomedical Research Institute of Salamanca (IBSAL), Salamanca, Spain

Julia Almeida & Alberto Orfao

274. Centre for Biomedical Network Research on Cancer (CIBERONC), Instituto de Salud Carlos III, Madrid, Spain

Julia Almeida, Alberto Orfao & Federico Rojo

275. Department of Genetics & Genomics, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital—Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Berta Almoguera, Carmen Ayuso, Marta Corton, Lidia Fernandez-Caballero, Ruth Fernández-Sánchez, Inés García, Rosario Lopez-Rodriguez, Andrea Martínez-Ramas, Laura Marzal, Pablo Minguez, Lorena Ondo & Cristina Villaverde

276. Human Genotyping—CEGEN Unit, Spanish National Cancer Research Centre, Madrid, Spain

María R. Alonso, Nuria Alvarez, Anna González-Neira, Belen Herraez, Rocio Nuñez-Torres & Guillermo Pita

277. Servicio de Medicina Interna, Hospital Universitario de Salamanca-IBSAL, Salamanca, Spain

Felipe Álvarez-Navia, Cristina Carbonell, Guillermo Hernández-Pérez, Amparo López-Bernús, Miguel Marcos & José-Ángel Martín-Oterino

278. Universidad de Salamanca, Salamanca, Spain

Felipe Álvarez-Navia, Moncef Belhassen-Garcia, Cristina Carbonell, Amparo López-Bernús, Miguel Marcos, José-Ángel Martín-Oterino & Pedro-Luis Sánchez

279. Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

Álvaro Andreu-Bernabeu, Celso Arango, Covadonga M. Diaz-Caneja, Javier González-Peñas & Mara Parellada

280. Clinical Pharmacology Service, Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

Maria Rosa Antonijoan & Maria Angeles Quijada

281. Biocruces Bizkai HRI, Barakaldo, Spain

Eunate Arana-Arri, Luis Castano, Ana B. de la Hoz, Aitor García-de-Vicuña, Natale Imaz-Ayo & Gustavo Perez-de-Nanclares

282. Cruces University Hospital, Osakidetza, Barakaldo, Spain

Eunate Arana-Arri

283. Hospital Infanta Elena, Madrid, Spain

Carlos Aranda, Maria Sanchez Carpintero, Carlos F. Castaño, Lidia Gagliardi, Mercedes García, Leticia García, María Herrera, Ángel Jiménez, María Rubio Olivera & Virginia Víctor

284. Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Carlos Aranda, Yolanda Cañadas, Maria Sanchez Carpintero, Carlos F. Castaño, Lidia Gagliardi, Mercedes García, Leticia García, María Herrera, Ángel Jiménez, María Rubio Olivera, Javier Ruiz-Hornillos & Virginia Víctor

285. Centre for Biomedical Network Research on Mental Health (CIBERSAM), Instituto de Salud Carlos III, Madrid, Spain

Celso Arango, Covadonga M. Diaz-Caneja, Javier González-Peñas & Mara Parellada

286. Fundación Hospital Infantil Universitario de San José, Bogotá, Colombia

Carolina Araque, César O. Enciso-Olivera & Hector D. Salamanca

287. Fundación Universitaria de Ciencias de la Salud, Bogotá, Colombia

Carolina Araque, Jessica G. Chaux, Walter G. Chaves-Santiago, César O. Enciso-Olivera, Mario Gómez-Duque, Luz D. Gutierrez-Castañeda, Luz Adriana Pinzón, Carlos S. Rivadeneira-

Chamorro, Marilyn Johanna Rodriguez, Hector D. Salamanca, John J. Sprockel & Lilian Torres-Tobar

288. Programa de Pós-graduação em Ciências da Saúde, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Nathalia K. Araujo, Marina S. Cruz, Raquel C. S. Dantas-Komatsu & Jeane F. P. Medeiros

289. Departamento de Medicina Clínica, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Izabel M. T. Araujo, Gabriela V. da Silva & Alice M. Duarte

290. Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, Universidade de Brasília, Brasilia, Brazil

Ana C. Arcanjo, Marcela C. Campos, Joana F. R. Nunes & Fabiola T. C. Silva

291. Colégio Marista de Brasilia, Brasilia, Brazil

Ana C. Arcanjo

292. Associação Brasileira de Educação e Cultura, Londrina, Brazil

Ana C. Arcanjo

293. Servicio de Medicina Interna, Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

Francisco Arnalich Fernández, José Ramon Arribas Lopez & Ana Méndez-Echevarria

294. Fundació Docència I Recerca Mutua Terrassa, Barcelona, Spain

María J. Arranz

295. Spanish National Cancer Research Center, CNIO Biobank, Madrid, Spain

Maria-Jesus Artiga & Eva Ortega-Paino

296. Departamento Patologia y Laboratorios, Fundación Santa Fe de Bogota, Bogotá, Colombia

Yubelly Avello-Malaver, Ana María Baldion, Viviana Barrera-Penagos, Oscar Martinez-Nieto, Diana Roa-Agudelo, Paula A. Rodriguez-Urrego & David A. Suarez-Zamora

297. Hospital General de Occidente, Zapopan, Mexico

Raúl C. Baptista-Rosas

298. Centro Universitario de Tonalá, Universidad de Guadalajara, Tonalá, Mexico

Raúl C. Baptista-Rosas

299. Centro de Investigación Multidisciplinario en Salud, Universidad de Guadalajara, Guadalajara, Mexico

Raúl C. Baptista-Rosas, Luis D. Hernandez-Ortega & Arieh R. Mercado-Sesma

300. Instituto Murciano de Investigación Biosanitaria (IMIB-Arrixaca), Murcia, Spain

María Barreda-Sánchez, Enrique Bernal, Josefina Garcia-García, Elisa García-Vázquez, Encarna Guillen-Navarro, M. Teresa Herranz, Rubén Jara, Antonio Moreno-Docón & M. Elena Pérez-Tomás

301. Universidad Católica San Antonio de Murcia (UCAM), Murcia, Spain

María Barreda-Sánchez

302. Servicio de Medicina Interna-Unidad de Enfermedades Infecciosas, Hospital Universitario de Salamanca-IBSAL, Salamanca, Spain

Moncef Belhassen-Garcia

303. Department of Internal Medicine, Hospital Universitario de Fuenlabrada, Madrid, Spain

David Bernal-Bello

304. Laboratorio de Vigilancia Molecular Aplicada, Escola Tecnica de Saúde, Pará, Brazil

Joao F. Bezerra

305. Genetics Postgraduate Program, Federal University of Pernambuco, Recife, Brazil

Marcos A. C. Bezerra

306. Servicio de Alergia, Hospital Universitario Infanta Leonor, Madrid, Spain

Natalia Blanca-López & Laura Martin-Pedraza

307. Servicio de Medicina Intensiva, Hospital Universitario del Tajo, Toledo, Spain

Rafael Blancas & Óscar Martínez-González

308. Hospital Universitario Mutua Terrassa, Barcelona, Spain

Lucía Boix-Palop & Anna Sangil

309. Servicio de Farmacología, Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

Alberto Borobia, Irene García-García & Alicia Marín Candon

310. Alcaldía de Barranquilla, Secretaría de Salud, Barranquilla, Colombia

Elsa Bravo, Humberto Mendoza Charris & Xenia Morelos-Arnedo

311. Xenética Cardiovascular, Instituto de Investigación Sanitaria de Santiago (IDIS), Santiago de Compostela, Spain

María Brion

312. Centre for Biomedical Network Research on Cardiovascular Diseases (CIBERCV), Instituto de Salud Carlos III, Madrid, Spain

María Brion, Ramón Brugada, Alexandra Pérez-Serra & Mel·lina Pinsach-Abuin

313. Unidad de Infección Viral e Inmunidad, Centro Nacional de Microbiología (CNM), Instituto de Salud Carlos III (ISCIII), Madrid, Spain

Óscar Brochado-Kith, Francisco C. Ceballos, Amanda Fernández-Rodríguez, María A. Jimenez-Sousa, María Martín-Vicente, Salvador Resino & Ana Virseda-Berdices

314. Cardiovascular Genetics Center, Institut d'Investigació Biomèdica Girona (IDIBGI), Girona, Spain

Ramón Brugada, Alexandra Pérez-Serra & Mel·lina Pinsach-Abuin

315. Medical Science Department, School of Medicine, University of Girona, Girona, Spain

Ramón Brugada

316. Cardiology Service, Hospital Josep Trueta, Girona, Spain

Ramón Brugada

317. Institute of Biomedicine of Seville (IBiS), Consejo Superior de Investigaciones Científicas (CSIC), University of Seville, Virgen del

Rocio University Hospital, Seville, Spain

Matilde Bustos & María A. Rodriguez-Hernandez

318. Division of Infectious Diseases, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Alfonso Cabello & Miguel Górgolas

319. Intensive Care Unit, Hospital Universitario Insular de Gran Canaria, Las Palmas de Gran Canaria, Spain

Juan J. Caceres-Agra

320. Hospital Universitario Mutua Terrassa, Terrassa, Spain

Esther Calbo, David Dalmau & Beatriz Dietl

321. Departemento de Medicina, Hospital Universitario Virgen del Rocío, Universidad de Sevilla, Seville, Spain

Enrique J. Calderón & Francisco J. Medrano

322. Centre for Biomedical Network Research on Epidemiology and Public Health (CIBERESP), Instituto de Salud Carlos III, Madrid, Spain

Enrique J. Calderón, Juan De la Cruz Troca, Vicente Friaza, Iolanda Jordan, Esther Lopez-Garcia, Vicente Martín, Francisco J. Medrano, Antonio J. J. Molina & Fernando Rodriguez-Artalejo

323. Instituto de Biomedicina de Sevilla, Seville, Spain

Enrique J. Calderón, Carmen de la Horra, Vicente Friaza, Francisco J. Medrano & Rubén Morilla

324. Facultad de Ciencias, Universidad de los Andes, Bogotá, Colombia

Shirley Camacho, María A. Castillo, María Claudia Lattig & Lorena Salazar-García

325. Department of Nutrition, University of Fortaleza (UNIFOR), Fortaleza, Brazil

Antonio Augusto F. Carioca

326. Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, São Paulo, Brazil

Thássia M. T. Carratto & Vitor M. S. Moraes

327. Andalusian Public Health System Biobank, Granada, Spain

José Antonio Carrillo-Avila & Sonia Panadero-Fajardo

328. Programa de Pós-Graduação em Ciências Farmacêuticas, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Maria C. C. Carvalho, Katiusse A. dos Santos & Karla S. C. Souza

329. Neuromuscular Unit, Neurology Department, Hospital Universitari de Bellvitge, L'Hospitalet de Llobregat, Barcelona, Spain

Carlos Casasnovas & Valentina Vélez-Santamaría

330. Bellvitge Biomedical Research Institute (IDIBELL), Neurometabolic Diseases Laboratory, L'Hospitalet de Llobregat, Barcelona, Spain

Carlos Casasnovas, Juan José Martínez, Laura Planas-Serra, Aurora Pujol, Agustí Rodriguez-Palmero, Montserrat Ruiz, Agatha Schlüter & Valentina Vélez-Santamaría

331. Osakidetza, Cruces University Hospital, Barakaldo, Spain

Luis Castano, Aitor García-de-Vicuña & Gustavo Perez-de-Nanclares

332. Centre for Biomedical Network Research on Diabetes and Metabolic Associated Diseases (CIBERDEM), Instituto de Salud Carlos III, Madrid, Spain

Luis Castano

333. University of Pais Vasco, UPV/EHU, Bizkaia, Spain

Luis Castano

334. Oncology and Genetics Unit, Instituto de Investigacion Sanitaria Galicia Sur, Xerencia de Xestion Integrada de Vigo-Servizo Galego de Saúde, Vigo, Spain

Jose E. Castelao

335. Hospital Universitario La Paz, Hospital Carlos III, Madrid, Spain

Aranzazu Castellano Candalija, Carmen Fernéndez Capitán, Cristina Marcelo Calvo, Alberto Moreno Fernández & Giorgina Gabriela Salgueiro Origlia

336. Hospital de San José, Sociedad de Cirugía de Bogota, Bogotá, Colombia

Walter G. Chaves-Santiago, Mario Gómez-Duque, Luz Adriana Pinzón & John J. Sprockel

337. Hospital Universitario Río Hortega, Valladolid, Spain

Rosa Conde-Vicente & Manuel Gonzalez-Sagrado

338. Servicio de Medicina Intensiva, Complejo Hospitalario Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

M. Lourdes Cordero-Lorenzana

339. Preventive Medicine Department, Valencia University, Valencia, Spain

Dolores Corella & Jose V. Sorlí

340. Centre for Biomedical Network Research on Physiopatology of Obesity and Nutrition (CIBEROBN), Instituto de Salud Carlos III, Madrid, Spain

Dolores Corella & Jose V. Sorlí

341. Department of Microgravity and Translational Regenerative Medicine, Otto von Guericke University, Magdeburg, Germany

Jose L. Cortes-Sanchez

342. Maternidade Escola Janário Cicco, Natal, Brazil

Tatiana X. Costa

343. Centro Nacional de Genotipado (CEGEN), Universidade de Santiago de Compostela, Santiago de Compostela, Spain

Raquel Cruz, María Gómez García, Inés Quintela, José Javier Suárez-Rama & Angel Carracedo

344. Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

Luisa Cuesta

345. Programa de Pós Graduação em Ciências da Saúde, Faculdade de Medicina, Universidade de Brasília, Brasilia, Brazil

Gabriela C. R. Cunha & Vanessa S. Souza

346. Fundació Docència I Recerca Mutua Terrassa, Terrassa, Spain

David Dalmau & Alex Serra-Llovich

347. Unidad de Genética, Hospital Universitario Mostoles, Madrid, Spain

M. Teresa Darnaude & Aranzazu Diaz de Bustamante

348. Internal Medicine Department, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Raimundo de Andrés

349. Universidade Federal do Rio Grande do Norte, Pós-graduação em Biotecnologia, Rede de Biotecnologia do Nordeste (Renorbio), Natal, Brazil

Jéssica N. G. de Araújo & Nayara S. Silva

350. Servicio de Medicina Interna, Hospital Universitario Severo Ochoa, Madrid, Spain

Carmen de Juan & Patricia Moreira-Escriche

351. Department of Preventive Medicine and Public Health, School of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

Juan De la Cruz Troca, Esther Lopez-Garcia & Fernando Rodriguez-Artalejo

352. IdiPaz (Instituto de Investigación Sanitaria Hospital Universitario La Paz), Madrid, Spain

Juan De la Cruz Troca, Esther Lopez-Garcia & Fernando Rodriguez-Artalejo

353. Instituto Aragonés de Ciencias de la Salud (IACS), Zaragoza, Spain

Alba De Martino-Rodríguez, Javier Gómez-Arrue, Beatriz González Álvarez, Delia Recalde & Carlos Tellería

354. Instituto Investigación Sanitaria Aragón (IIS-Aragon), Zaragoza, Spain

Alba De Martino-Rodríguez, Javier Gómez-Arrue, Beatriz González Álvarez, Delia Recalde & Carlos Tellería

355. Programa de Pós Graduação em Nutrição, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Julianna Lys de Sousa Alves Neri

356. Preventive Medicine Department, Instituto de Investigacion Sanitaria Galicia Sur, Xerencia de Xestion Integrada de Vigo-Servizo Galego de Saúde, Vigo, Spain

Victor del Campo-Pérez

357. Servicio de Medicina Interna, Hospital Universitario Virgen del Rocío, Seville, Spain

Juan Delgado-Cuesta

358. Departamento de Infectologia, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Manoella do Monte Alves

359. Hospital de Doenças Infecciosas Giselda Trigueiro, Natal, Brazil

Manoella do Monte Alves

360. Unidad Diagnóstico Molecular, Fundación Rioja Salud, La Rioja, Spain

Elena Domínguez-Garrido

361. Hospital Universitario Quironsalud Madrid, Madrid, Spain

Jose Echave-Sustaeta & Pablo Guisado-Vasco

362. Servicio de Cardiología, Hospital Universitario de Salamanca-IBSAL, Salamanca, Spain

Rocío Eiros, Pedro-Luis Sánchez & Clara Sánchez-Pablo

363. Servicio de Medicina Interna, Hospital Universitario Puerta de Hierro, Majadahonda, Spain

Gabriela Escudero

364. Biocruces Bizkaia Health Research Institute, Galdakao University Hospital, Osakidetza, Bizkaia, Spain

Pedro Pablo España & Carmen Mar

365. Instituto Regional de Investigación en Salud-Universidad Nacional de Caaguazú, Caaguazú, Paraguay

Gladys Mercedes Estigarribia Sanabria

366. Núcleo de Pesquisas em Oncologia, Universidade Federal do Pará, Belém, Brazil

Marianne R. Fernandes & Ney P. C. Santos

367. Departamento de Ensino e Pesquisa, Hospital Ophir Loyola, Belém, Brazil

Marianne R. Fernandes

368. Fundación Asilo San Jose, Santander, Spain

Ramón Fernández

369. Unidad de Enfermedades Infecciosas, Servicio de Medicina Interna, Hospital Universitario Puerta de Hierro, Instituto de Investigación Sanitaria Puerta de Hierro—Segovia de Arana, Madrid, Spain

Ana Fernández-Cruz

370. Universidad Nacional de Asunción, Facultad de Politécnica, Paraguay

María J. Fernandez-Nestosa

371. Urgencias Hospitalarias, Complejo Hospitalario Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

Uxía Fernández-Robelo

372. Grupo de Investigación en Interacciones Gen-Ambiente y Salud (GIIGAS), Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain

Tania Fernández-Villa

373. Pediatrics Department, Hospital Universitario Niño Jesús, Madrid, Spain

Patricia Flores-Pérez

374. Unitat de Malalties Infeccioses i Importades, Servei de Pediatría, Infectious and Imported Diseases, Pediatric Unit, Hospital Universitari Sant Joan de Deú, Barcelona, Spain

Victoria Fumadó

375. Microbiology Department, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospita, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Ignacio Gadea & Celia Perales

376. Hospital de Niños Ricardo Gutierrez, Buenos Aires, Argentina

Cristina Galoppo, Angela Gentile, Florencia Guaragna, Perez Maria Jazmin, Ailen Lauriente, Pablo Neira, Virginia Olivar, German Ezequiel Rodriguez Novoa & Alejandro Teper

377. University of Salamanca, Biomedical Research Institute of Salamanca (IBSAL), Salamanca, Spain

Andrés C. García-Montero

378. Department of Immunology, IRYCIS, Hospital Universitario Ramón y Cajal, Madrid, Spain

Ana García-Soidán, Israel Nieto-Gañán & Roberto Pariente-Rodríguez

379. Servicio de Medicina Intensiva, Hospital Infanta Elena, Madrid, Spain

María Carmen García Torrejón

380. Servicio de Genética, Hospital Universitario de Getafe, Madrid, Spain

Belén Gil-Fournier & Soraya Ramiro León

381. Pneumology Department, Hospital General Universitario Gregorio Marañón (iiSGM), Madrid, Spain

Ángela Gómez Sacristán

382. Ministerio de Salud Ciudad de Buenos Aires, Buenos Aires, Argentina

Fernan Gonzalez Bernaldo de Quirós

383. Unidad de Apoyo a la Investigación, Hospital Clinico Universitario de Valladolid, Valladolid, Spain

Hugo Gonzalo Benito & Pedro Martinez-Paz

384. Departamento de Cirugía, Universidad de Valladolid, Valladolid, Spain

Oscar Gorgojo-Galindo & Eduardo Tamayo

385. Secretaria Municipal de Saude de Apodi, Natal, Brazil

Genilson P. Guegel

386. Sección Genética Médica, Servicio de Pediatría, Hospital Clínico Universitario Virgen de la Arrixaca, Servicio Murciano de Salud,

Murcia, Spain

Encarna Guillen-Navarro

387. Departamento Cirugía, Pediatría, Obstetricia y Ginecología, Facultad de Medicina, Universidad de Murcia (UMU), Murcia, Spain

Encarna Guillen-Navarro

388. Grupo Clínico Vinculado, Centre for Biomedical Network Research on Rare Diseases (CIBERER), Instituto de Salud Carlos III, Madrid, Spain

Encarna Guillen-Navarro

389. Servicio de Análisis Clínicos e Inmunología, Hospital Universitario Virgen de las Nieves, Granada, Spain

Juan F. Gutiérrez-Bautista, Pilar Jiménez, Antonio Rodriguez-Nicolas & Francisco Ruiz-Cabello

390. Hospital Universitario Centro Dermatológico Federico Lleras Acosta, Bogotá, Colombia

Luz D. Gutierrez-Castañeda

391. Intermediate Respiratory Care Unit, Department of Pneumology, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Sarah Heili-Frades

392. Clinica Comfamiliar Risaralda, Pereira, Colombia

Estefania Hernandez, Ricardo Martínez, Juan José Montoya, Manuel Pacheco & Gloria L. Porras-Hurtado

393. Centro Universitario de Tonalá, Universidad de Guadalajara, Guadalajara, Mexico

Luis D. Hernandez-Ortega & Arieh R. Mercado-Sesma

394. Unidad de Cuidados Intensivos, Hospital Clínico Universitario de Santiago (CHUS), Sistema Galego de Saúde (SERGAS), Santiago de Compostela, Spain

Rebeca Hernández-Vaquero, Pedro Rascado Sedes, Montserrat Robelo Pardo & Emilio Rodríguez-Ruiz

395. Plataforma de Farmacogenética, IIS La Fe, Valencia, Spain

María José Herrero

396. Departamento de Farmacología, Universidad de Valencia, Valencia, Spain

María José Herrero

397. Data Analysis Department, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Antonio Herrero-Gonzalez & Juan Carlos Taracido-Fernandez

398. Infectious Diseases Service, Hospital del Mar, Barcelona, Spain

Juan P. Horcajada

399. Institut Hospital del Mar d'Investigacions Mèdiques (IMIM), Barcelona, Spain

Juan P. Horcajada

400. CEXS-Universitat Pompeu Fabra, Spanish Network for Research in Infectious Diseases (REIPI), Barcelona, Spain

Juan P. Horcajada

401. Biocruces Bizkaia Health Research Institute, Basurto University Hospital, Osakidetza, Bizkaia, Spain

Maider Intxausti-Urrutibeaskoa & Cristina Sancho-Sainz

402. Infectious Diseases, Microbiota and Metabolism Unit, Center for Biomedical Research of La Rioja (CIBIR), Logroño, Spain

María Íñiguez, José A. Oteo, Patricia Pérez-Matute, Emma Recio-Fernández & Pablo Villoslada-Blanco

403. Sabin Medicina Diagnóstica, São Paulo, Brazil

Rafael H. Jacomo

404. Opthalmology Department, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Ignacio Jiménez-Alfaro

405. Pediatric Critical Care Unit, Hospital Sant Joan de Deu, Barcelona, Spain

Iolanda Jordan

406. Paediatric Intensive Care Unit, Agrupación Hospitalaria Clínic-Sant Joan de Déu, Esplugues de Llobregat, Barcelona, Spain

Iolanda Jordan

407. Department of Immunology, Hospital Universitario 12 de Octubre, Madrid, Spain

Rocío Laguna-Goya, María Lasa-Lazaro, Esther Mancebo & Estela Paz-Artal

408. Transplant Immunology and Immunodeficiencies Group, Instituto de Investigación Sanitaria Hospital 12 de Octubre (imas12), Madrid, Spain

Rocío Laguna-Goya, María Lasa-Lazaro, Esther Mancebo & Estela Paz-Artal

409. SIGEN Alianza Universidad de los Andes, Fundación Santa Fe de Bogotá, Bogotá, Colombia

María Claudia Lattig & Oscar Martinez-Nieto

410. Medicina Intensiva, Hospital General de Segovia, Segovia, Spain

Anabel Liger Borja, María Lozano-Espinosa & Caridad Martín-López

411. Clinical Trials Unit, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Lucía Llanos

412. IMDEA-Food Institute, CEI UAM+CSIC, Madrid, Spain

Esther Lopez-Garcia & Fernando Rodriguez-Artalejo

413. Servicio de Enfermedades Infecciosas, Hospital Universitario Virgen de las Nieves, Granada, Spain

Miguel A. López-Ruz

414. Instituto de Investigación Biosanitaria de Granada (ibs.GRANADA), Granada, Spain

Miguel A. López-Ruz & Francisco Ruiz-Cabello

415. Departamento de Medicina, Universidad de Granada, Granada, Spain

Miguel A. López-Ruz

416. Servicio de Inmunología, Hospital Universitario La Paz-IDIPAZ, Madrid, Spain

Eduardo López Granados

417. Lymphocyte Pathophysiology in Immunodeficiencies Group, La Paz Institute for Health Research (IdiPAZ), Madrid, Spain

Eduardo López Granados

418. Intensive Care Unit, Hospital Universitario de Canarias, La Laguna, Spain

Leonardo Lorente

419. Dirección General de Salud Pública, Consejería de Sanidad, Junta de Castilla y León, Valladolid, Spain

José E. Lozano

420. Departamento de Analises Clinicas e Toxicologicas, Universidade Federal do Rio Grande do Norte, Natal, Brazil

Andre D. Luchessi & Vivian N. Silbiger

421. Epidemiology, Fundación Jiménez Díaz, Madrid, Spain

Ignacio Mahillo

422. Department of Medicine, Universidad Autónoma de Madrid, Madrid, Spain

Ignacio Mahillo

423. Instituto de Biomedicina (IBIOMED), Universidad de León, León, Spain

Alba Marcos-Delgado, Vicente Martín & Antonio J. J. Molina

424. Intensive Care Unit, Hospital Universitario N. S. de Candelaria, Santa Cruz de Tenerife, Spain

María M. Martín

425. Preventive Medicine Department, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

María Dolores Martín

426. Departamento de Medicina, Universidad de Valladolid, Valladolid, Spain

Marta Martin-Fernandez

427. Servicio de Medicina Intensiva, Hospital Universitario Infanta Leonor, Madrid, Spain

Amalia Martinez

428. Servicio de Medicina Interna, Sanatorio Franchin, Buenos Aires, Argentina

Eleno Martínez-Aquino

429. Unidad de Genética y Genómica Islas Baleares, Hospital Universitario Son Espases, Islas Baleares, Spain

Iciar Martinez-Lopez

430. Unidad de Diagnóstico Molecular y Genética Clínica, Hospital Universitario Son Espases, Islas Baleares, Spain

Iciar Martinez-Lopez

431. Genomics of Complex Diseases Unit, Research Institute of Hospital de la Santa Creu i Sant Pau, IIB Sant Pau, Barcelona, Spain

Angel Martinez-Perez & José Manuel Soria

432. Faculdade de Medicina, Universidade de Brasília, Brasilia, Brazil

Juliana F. Mazzeu, Silviene F. Oliveira & Renata R. Sousa

433. Programa de Pós-Graduação em Ciências Médicas, Universidade de Brasília, Brasilia, Brazil

Juliana F. Mazzeu & Silviene F. Oliveira

434. Programa de Pós-Graduação em Ciências da Saúde, Universidade de Brasília, Brasilia, Brazil

Juliana F. Mazzeu & Silviene F. Oliveira

435. Hospital das Forças Armadas, Brasília, Brazil

Kelliane A. Medeiros, Susana M. T. Pinho & Adriana P. Ribeiro

436. Exército Brasileiro, Cruzeiro, Brazil

Kelliane A. Medeiros & Adriana P. Ribeiro

437. Hospital El Bierzo, Gerencia de Asistencia Sanitaria del Bierzo (GASBI), Gerencia Regional de Salud (SACYL), Ponferrada, Spain

Xose M. Meijome

438. Grupo INVESTEN, Instituto de Salud Carlos III, Madrid, Spain

Xose M. Meijome

439. Unidad de Cuidados Intensivos, Complejo Universitario de A Coruña (CHUAC), Sistema Galego de Saúde (SERGAS), A Coruña, Spain

Natalia Mejuto-Montero & Xiana Taboada-Fraga

440. Unidad Cuidados Intensivos, Hospital El Bierzo, León, Spain

Eleuterio Merayo Macías

441. Familial Cancer Clinical Unit, Spanish National Cancer Research Centre, Madrid, Spain

Fátima Mercadillo & Miguel Urioste

442. Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos (HCSC), Madrid, Spain

Elena Molina-Roldán

443. Departamento de Enfermería, Universidad de Sevilla, Seville, Spain

Rubén Morilla

444. Hospital General Universitario Gregorio Marañón (IiSGM), Madrid, Spain

Patricia Muñoz García

445. ERN-ITHACA-European Reference Network on Rare Congenital Malformations and Rare Intellectual Disability, Brussels, Belgium

Julian Nevado, Jair Antonio Tenorio Castaño & Pablo Lapunzina

446. Unidad de Genética y Genómica Islas Baleares, Unidad de Diagnóstico Molecular y Genética Clínica, Hospital Universitario Son Espases, Islas Baleares, Spain

Antònia Obrador-Hevia

447. Instituto de Investigación Sanitaria Islas Baleares (IdISBa), Islas Baleares, Spain

Antònia Obrador-Hevia

448. Programa de Pós-Graduação em Biologia Animal, Universidade de Brasília, Brasília, Brazil

Silviene F. Oliveira

449. Programa de Pós-Graduação Profissional em Ensino de Biologia, Universidade de Brasília, Brasília, Brazil

Silviene F. Oliveira

450. Anatomía Patológica, Instituto de Investigación Sanitaria San Carlos (IdISSC), Hospital Clínico San Carlos (HCSC), Madrid, Spain

Luis Ortega

451. Tecnológico de Monterrey, Monterrey, Mexico

Rocio Ortiz-Lopez

452. Centro de Investigación en Anomalías Congénitas y Enfermedades Raras (CIACER), Universidad Icesi, Cali, Colombia

Harry Pachajoa

453. Departamento de Genetica, Fundación Valle del Lili, Cali, Colombia

Harry Pachajoa

454. Department of Immunology, Ophthalmology and ENT, Universidad Complutense de Madrid, Madrid, Spain

Estela Paz-Artal

455. Department of Neumology, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Germán Peces-Barba & Felipe Villar

456. Hospital Nuestra Señora de Sonsoles, Ávila, Spain

Miguel S. Pedromingo Kus & Ronald P. Torres Gutiérrez

457. Inditex, A Coruña, Spain

Patricia Perez

458. Intensive Care Department, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

César Pérez & Arnoldo Santos

459. Servicio de Microbiología Clínica, Hospital Universitario Príncipe de Asturias, Madrid, Spain

Felipe Pérez-García

460. Departamento de Biomedicina y Biotecnología, Facultad de Medicina y Ciencias de la Salud, Universidad de Alcalá de Henares, Madrid, Spain

Felipe Pérez-García

461. GENYCA, Madrid, Spain

Teresa Perucho & Eva Ruiz-Casares

462. Marinha do Brasil, Brasil, Brazil

Susana M. T. Pinho

463. Universidade de Brasília, Brasilia, Brazil

Susana M. T. Pinho & Adriana P. Ribeiro

464. Neuromuscular Diseases Unit, Department of Neurology, Hospital de la Santa Creu i Sant Pau, Universitat Autònoma de Barcelona, Barcelona, Spain

Francesc Pla-Junca & Sonia Segovia

465. Unidad de Investigación Médica en Enfermedades Infecciosas y Parasitarias, Instituto Mexicano del Seguro Social (IMSS), Centro Médico Nacional Siglo XXI, Mexico City, Mexico

Ericka N. Pompa-Mera

466. Catalan Institution of Research and Advanced Studies (ICREA), Barcelona, Spain

Aurora Pujol

467. Drug Research Centre, Institut d'Investigació Biomèdica Sant Pau, IIB-Sant Pau, Barcelona, Spain

Maria Angeles Quijada

468. Departamento de Genetica, Clinica imbanaco, Cali, Colombia

Diana Ramirez-Montaño

469. Department of Immunology, Hospital Universitario de Gran Canaria Dr Negrín, Las Palmas de Gran Canaria, Spain

Carlos Rodriguez-Gallego

470. Pediatrics Department, University Hospital Germans Trias i Pujol, Badalona, Spain

Agustí Rodriguez-Palmero

471. Department of Pathology, Biobank, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Federico Rojo & Sandra Zazo

472. Faculdade de Ciências da Saúde, Universidade de Brasília, Brasilia, Brazil

Lidia S. Rosa

473. Servicio de Medicina Interna, Hospital Universitario Virgen de las Nieves, Granada, Spain

Antonio Rosales-Castillo

474. Grupo de Ciencias Básicas en Salud (CBS), Fundación Universitaria de Ciencias de la Salud, Bogotá, Colombia

Cladelis Rubio

475. Sociedad de Cirugía de Bogotá, Hospital de San José, Bogotá, Colombia

Cladelis Rubio

476. Departamento Bioquímica, Biología Molecular e Inmunología III, Universidad de Granada, Granada, Spain

Francisco Ruiz-Cabello

477. Allergy Unit, Hospital Infanta Elena, Madrid, Spain

Javier Ruiz-Hornillos

478. Faculty of Medicine, Universidad Francisco de Vitoria, Madrid, Spain

Javier Ruiz-Hornillos

479. Hospital Universitario Infanta Leonor, Madrid, Spain

Pablo Ryan & Jesús Troya

480. Complutense University of Madrid, Madrid, Spain

Pablo Ryan

481. Gregorio Marañón Health Research Institute (IiSGM), Madrid, Spain

Pablo Ryan

482. Reumathology Service, Instituto de Investigación Sanitaria–Fundación Jiménez Díaz University Hospital, Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain

Olga Sánchez-Pernaute

483. Biobank, Puerta de Hierro-Segovia de Arana Health Research Institute, Madrid, Spain

Antonio J. Sánchez López

484. Universidad Rey Juan Carlos, Madrid, Spain

Jorge Sánchez Redondo

485. The John Walton Muscular Dystrophy Research Centre, Newcastle University and Newcastle Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK

Sonia Segovia

486. Neuromuscular Unit, Neuropediatrics Department, Institut de Recerca Sant Joan de Déu, Hospital Sant Joan de Déu, Barcelona, Spain

Sonia Segovia

487. Casa de Saúde São Lucas, Natal, Brazil

Miguel A. Sicolo

488. Hospital Rio Grande, Natal, Brazil

Miguel A. Sicolo

489. Intensive Care Unit, Hospital Universitario de Gran Canaria Dr Negrín, Las Palmas de Gran Canaria, Spain

Jordi Solé-Violán

490. Universidad Fernando Pessoa Canarias, Las Palmas de Gran Canaria, Spain

Jordi Solé-Violán

491. Servicio de Anestesiologia y Reanimación, Hospital Clinico Universitario de Valladolid, Valladolid, Spain

Eduardo Tamayo

492. Servicio de Hematologia y Hemoterapia, Hospital Clinico Universitario de Valladolid, Valladolid, Spain

Alvaro Tamayo-Velasco

493. Hospital Universitario Lauro Wanderley, João Pessoa, Brazil

Nathali A. C. Tavares & Romero H. T. Vasconcelos

494. Servicio de Medicina Interna, Hospital Universitario Infanta Leonor, Madrid, Spain

Juan Torres-Macho

495. University Hospital of Burgos, Burgos, Spain

Juan Valencia-Ramos

496. Universidad de Sevilla, Seville, Spain

Agustín Valido

497. Fundación Santa Fe de Bogota, Instituto de Servicios Medicos de Emergencia y Trauma, Bogotá, Colombia

Juan Pablo Vargas Gallo

498. Universidad de los Andes, Bogotá, Colombia

Juan Pablo Vargas Gallo

499. Quironprevención, A Coruña, Spain

Belén Varón

500. Junta de Castilla y León, Consejería de Sanidad, Valladolid, Spain

Tomas Vega

501. Gerencia Atención Primaria de Burgos, Burgos, Spain

Santiago Velasco-Quirce

502. Immunogenetics–Histocompatibility group, Servicio de Inmunología, Instituto de Investigación Sanitaria Puerta de Hierro, Segovia de Arana, Madrid, Spain

Carlos Vilches

503. Department of Infectious Diseases, Hospital del Mar, Barcelona, Spain

Judit Villar-Garcia

504. IMIM—Hospital del Mar Medical Research Institute, Institut Hospital del Mar d'Investigacions Mediques, Barcelona, Spain

Judit Villar-Garcia

505. Department of Medicine, Universitat Autònoma de Barcelona, Barcelona, Spain

Judit Villar-Garcia

506. Consejería de Sanidad, Comunidad de Madrid, Madrid, Spain

Antonio Zapatero-Gaviria

507. Centro para el Desarrollo de la Investigación Científica, Asunción, Paraguay

Ruth Zarate

508. School of Informatics, University of Edinburgh, Edinburgh, UK

Beatrice Alex, Benjamin Bach & James Scott-Brown

509. Section of Biomolecular Medicine, Division of Systems Medicine, Department of Metabolism, Digestion and Reproduction, Sir Alexander Fleming Building, Imperial College London, London, UK

Petros Andrikopoulos, Kanta Chechi, Marc-Emmanuel Dumas, Julian Griffin, Sonia Liggi, Michael Olanipekun, Anthonia Osagie & Zoltan Takats

510. Section of Genomic and Environmental Medicine, Respiratory Division, National Heart and Lung Institute, London, UK

Petros Andrikopoulos, Marc-Emmanuel Dumas, Michael Olanipekun & Anthonia Osagie

511. Section of Molecular Virology, Imperial College London, London, UK

Wendy S. Barclay

512. Antimicrobial Resistance and Hospital Acquired Infection Department, Public Health England, London, UK

Meera Chand

513. Department of Epidemiology and Biostatistics, School of Public Health, Faculty of Medicine, Imperial College London, London, UK

Kanta Chechi

514. Department of Infectious Disease, Imperial College London, London, UK

Graham S. Cooke & Shiranee Sriskandan

515. MRC-University of Glasgow Centre for Virus Research, Glasgow, UK

Ana da Silva Filipe, Antonia Y. W. Ho, Massimo Palmarini, David L. Robertson, Janet T. Scott & Emma C. Thomson

516. The Florey Institute for Host-Pathogen Interactions, Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK

Thushan de Silva

517. Centre for Medical Informatics, The Usher Institute, University of Edinburgh, Edinburgh, UK

Annemarie B. Docherty, Ewen M. Harrison, Thomas M. Drake, Cameron J. Fairfield, Stephen R. Knight, Kenneth A. Mclean, Derek Murphy, Lisa Norman, Riinu Pius & Catherine A. Shaw

518. National Phenome Centre, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK

Gonçalo dos Santos Correia, Matthew Lewis, Lynn Maslen, Caroline Sands & Panteleimon Takis

519. Section of Bioanalytical Chemistry, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK

Gonçalo dos Santos Correia, Matthew Lewis, Lynn Maslen, Caroline Sands & Panteleimon Takis

520. European Genomic Institute for Diabetes, CNRS UMR 8199, INSERM UMR 1283, Institut Pasteur de Lille, Lille University Hospital, University of Lille, Lille, France

Marc-Emmanuel Dumas

521. McGill University and Genome Quebec Innovation Centre, Montréal, Quebec, Canada

Marc-Emmanuel Dumas

522. National Infection Service, Public Health England, London, UK

Jake Dunning & Maria Zambon

523. Liverpool School of Tropical Medicine, Liverpool, UK

Tom Fletcher

524. Institute of Microbiology and Infection, University of Birmingham, Birmingham, UK

Christopher A. Green

525. Department of Molecular and Clinical Cancer Medicine, University of Liverpool, Liverpool, UK

William Greenhalf

526. Institute for Global Health, University College London, London, UK

Rishi K. Gupta

527. Department of Infectious Diseases, Queen Elizabeth University Hospital, Glasgow, UK

Antonia Y. W. Ho

528. University of Liverpool, Liverpool, UK

Karl Holden

529. Virology Reference Department, National Infection Service, Public Health England, London, UK

Samreen Ijaz

530. Department of Pharmacology, University of Liverpool, Liverpool, UK

Say Khoo

531. Nuffield Department of Medicine, Peter Medawar Building for Pathogen Research, University of Oxford, Oxford, UK

Paul Klenerman

532. Translational Gastroenterology Unit, Nuffield Department of Medicine, University of Oxford, Oxford, UK

Paul Klenerman

533. Nottingham University Hospitals NHS Trust, Nottingham, UK

Wei Shen Lim

534. Nuffield Department of Medicine, John Radcliffe Hospital, Oxford, UK

Alexander J. Mentzer

535. Department of Microbiology/Infectious Diseases, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford, UK

Alexander J. Mentzer

536. ISARIC Global Support Centre, Centre for Tropical Medicine and Global Health, Nuffield Department of Medicine, University of Oxford, Oxford, UK

Laura Merson, Louise Sigfrid & Gail Carson

537. Institute of Infection, Veterinary and Ecological Sciences, University of Liverpool, Liverpool, UK

Shona C. Moore, William A. Paxton & Georgios Pollakis

538. Division of Infection and Immunity, University College London, London, UK

Mahdad Noursadeghi

539. Molecular and Clinical Cancer Medicine, Institute of Systems, Molecular and Integrative Biology, University of Liverpool, Liverpool, UK

Carlo Palmieri

540. Clatterbridge Cancer Centre NHS Foundation Trust, Liverpool, UK

Carlo Palmieri

541. NIHR Health Protection Research Unit in Emerging and Zoonotic Infections, Liverpool, UK

William A. Paxton & Georgios Pollakis

542. Centre for Clinical Infection and Diagnostics Research, Department of Infectious Diseases, School of Immunology and Microbial Sciences, King’s College London, London, UK

Nicholas Price

543. Department of Infectious Diseases, Guy’s and St Thomas’ NHS Foundation Trust, London, UK

Nicholas Price

544. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, UK

Andrew Rambaut

545. Department of Pediatrics and Virology, St Mary's Medical School Building, Imperial College London, London, UK

Vanessa Sancho-Shimizu

546. NHS Greater Glasgow & Clyde, Glasgow, UK

Janet T. Scott

547. Walton Centre NHS Foundation Trust, Liverpool, UK

Tom Solomon

548. MRC Centre for Molecular Bacteriology and Infection, Imperial College London, London, UK

Shiranee Sriskandan

549. Department of Child Life and Health, University of Edinburgh, Edinburgh, UK

Olivia V. Swann

550. National Phenome Centre, Division of Systems Medicine, Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK

Zoltan Takats

551. Blood Borne Virus Unit, Virus Reference Department, National Infection Service, Public Health England, London, UK

Richard S. Tedder

552. Transfusion Microbiology, National Health Service Blood and Transplant, London, UK

Richard S. Tedder

553. Department of Medicine, Imperial College London, London, UK

Richard S. Tedder

554. Department of Infection, Immunity and Cardiovascular Disease, University of Sheffield, Sheffield, UK

A. A. Roger Thompson

555. Tropical & Infectious Disease Unit, Royal Liverpool University Hospital, Liverpool, UK

Lance C. W. Turtle

556. Liverpool Clinical Trials Centre, University of Liverpool, Liverpool, UK

Marie Connor, Jo Dalton, Carrol Gamble, Michelle Girvan, Sophie Halpin, Janet Harrison, Clare Jackson, Laura Marsh, Stephanie Roberts, Egle Saviciute & Chloe Donohue

557. Public Health Scotland, Edinburgh, UK

Susan Knight, Eva Lahnsteiner & Sarah Tait

558. Centre for Health Informatics, Division of Informatics, Imaging and Data Science, School of Health Sciences, Faculty of Biology, Medicine and Health, University of Manchester, Manchester Academic Health Science Centre, Manchester, UK

Gary Leeming

559. EPCC, University of Edinburgh, Edinburgh, UK

Lucy Norris

560. ISARIC, Global Support Centre, COVID-19 Clinical Research Resources, Epidemic diseases Research Group, Oxford (ERGO), University of Oxford, Oxford, UK

James Lee & Daniel Plotkin

561. Institute of Infection, Veterinary and Ecological Sciences, Faculty of Health and Life Sciences, University of Liverpool, Liverpool, UK

Cara Donegan & Rebecca G. Spencer

562. 23andMe, Sunnyvale, CA, USA

Janie F. Shelton, Anjali J. Shastri, Chelsea Ye, Catherine H. Weldon, Teresa Filshtein-Sonmez, Daniella Coker, Antony Symons, Stella Aslibekyan & Adam Auton

563. Human genetics R&D, GSK Medicines Research Centre, Target Sciences R&D, Stevenage, UK

Jorge Esparza-Gordillo

Consortia

GenOMICC Investigators

- **Co-Investigators**
    - J. Kenneth Baillie
    - , Colin Begg
    - , Sara Clohisey
    - , Charles Hinds
    - , Peter Horby
    - , Julian Knight
    - , Lowell Ling
    - , David Maslove
    - , Danny McAuley
    - , Johnny Millar
    - , Hugh Montgomery
    - , Alistair Nichol
    - , Peter J. M. Openshaw

- , Alexandre C. Pereira
- , Chris P. Ponting
- , Kathy Rowan
- , Malcolm G. Semple
- , Manu Shankar-Hari
- , Charlotte Summers
- & Timothy Walsh

- **Management, Laboratory and Data Team**

  - Emma Aitkin
  - , Latha Aravindan
  - , Ruth Armstrong
  - , J. Kenneth Baillie
  - , Heather Biggs
  - , Ceilia Boz
  - , Adam Brown
  - , Primmy Chikowore
  - , Richard Clark
  - , Sara Clohisey
  - , Audrey Coutts
  - , Judy Coyle
  - , Louise Cullum
  - , Sukamal Das
  - , Nicky Day
  - , Lorna Donnelly
  - , Esther Duncan
  - , Angie Fawkes
  - , Paul Finernan
  - , Max Head Fourman
  - , Anita Furlong
  - , James Furniss
  - , Bernadette Gallagher
  - , Tammy Gilchrist
  - , Ailsa Golightly
  - , Fiona Griffiths
  - , Katarzyna Hafezi

- , Debbie Hamilton
- , Ross Hendry
- , Naomi Kearns
- , Andy Law
- , Dawn Law
- , Rachel Law
- , Sarah Law
- , Rebecca Lidstone-Scott
- , Christen Lauder
- , Louise Macgillivray
- , Alan Maclean
- , Hanning Mal
- , Sarah McCafferty
- , Ellie McMaster
- , Jen Meikle
- , Shona C. Moore
- , Kirstie Morrice
- , Lee Murphy
- , Sheena Murphy
- , Hellen Mybaya
- , Miranda Odam
- , Wilna Oosthuyzen
- , Chenqing Zheng
- , Jiantao Chen
- , Nick Parkinson
- , Trevor Paterson
- , Petra Tucker
- , Katherine Schon
- , Andrew Stenhouse
- , Mihaela Das
- , Maaike Swets
- , Helen Szoor-McElhinney
- , Filip Taneski
- , Lance Turtle
- , Tony Wackett
- , Mairi Ward
- , Jane Weaver

    - , Nicola Wrobel
    - & Marie Zechner

- **Guys and St Thomas' Hospital, London, UK**

    - Jacqueline Pan
    - , Neus Grau
    - , Tim Owen Jones
    - , Rosario Lim
    - , Martina Marotti
    - , Christopher Whitton
    - , Aneta Bociek
    - , Sara Campos
    - , Gill Arbane
    - , Manu Shankar-Hari
    - , Marlies Ostermann
    - , Mina Cha
    - , Fabiola DAmato
    - , Eirini Kosifidou
    - , Shelley Lorah
    - & Kyma Morera

- **James Cook University Hospital, Middlesbrough, UK**

    - Laura Brady
    - , Keith Hugill
    - , Jeremy Henning
    - , Stephen Bonner
    - , Evie Headlam
    - , Jessica Jones
    - , Abigail List
    - , Joanne Morley
    - , Amy Welford
    - , Bobette Kamangu
    - , Anitha Ratnakumar
    - & Abiola Shoremekun

- **Barts Health NHS Trust, London, UK**
  - Zoe Alldis
  - , Raine Astin-Chamberlain
  - , Fatima Bibi
  - , Jack Biddle
  - , Sarah Blow
  - , Matthew Bolton
  - , Catherine Borra
  - , Ruth Bowles
  - , Maudrian Burton
  - , Yasmin Choudhury
  - , Amber Cox
  - , Amy Easthope
  - , Patrizia Ebano
  - , Stavros Fotiadis
  - , Jana Gurasashvili
  - , Rosslyn Halls
  - , Pippa Hartridge
  - , Delordson Kallon
  - , Jamila Kassam
  - , Ivone Lancoma-Malcolm
  - , Maninderpal Matharu
  - , Peter May
  - , Oliver Mitchelmore
  - , Tabitha Newman
  - , Mital Patel
  - , Jane Pheby
  - , Irene Pinzuti
  - , Zoe Prime
  - , Oleksandra Prysyazhna
  - , Julian Shiel
  - , Melanie Taylor
  - , Carey Tierney
  - , Olivier Zongo
  - , Suzanne Wood
  - , Anne Zak

- & David Collier

- **Royal Stoke University Hospital, Staffordshire, UK**
  - Manuela Mundy
  - , Christopher Thompson
  - , Lisa Pritchard
  - , Minnie Gellamucho
  - , David Cartlidge
  - , Nageswar Bandla
  - , Lucy Bailey
  - , Michelle Davies
  - , Jane Delaney
  - & Leanne Scott

- **North Middlesex University Hospital NHS trust, London, UK**
  - Marwa Abdelrazik
  - , Frater Alasdair
  - , David Carter
  - , Munzir Elhassan
  - , Arunkumar Ganesan
  - , Samuel Jenkins
  - , Zoe Lamond
  - , Dharam Purohit
  - , Kumar Rohit
  - , Malik Saleem
  - , Alanna Wall
  - , Kugan Xavier
  - , Dhanalaksmi Bakthavatsalam
  - , Kirolos Gehad
  - , Pakeerathan Gnanapragasam
  - , Kapil Jain
  - , Swati Jain
  - , Abdul Malik
  - , Naveen Pappachan

- , Jeronimo Moreno-Cuesta
- , Anne Haldeos
- , Rachel Vincent
- & Maryjane Oziegb

- **King's College Hospital, London, UK**

  - Anna Cavazza
  - , Maeve Cockrell
  - , Eleanor Corcoran
  - , Maria Depante
  - , Clare Finney
  - , Ellen Jerome
  - , Abigail Knighton
  - , Monalisa Nayak
  - , Evita Pappa
  - , Rohit Saha
  - , Sian Saha
  - , Andrew Dodd
  - , Kevin O'Reilly
  - , Mark McPhail
  - , Emma Clarey
  - , Harriet Noble
  - & John Smith

- **Charing Cross Hospital, St Mary's Hospital and Hammersmith Hospital, London, UK**

  - Phoebe Coghlan
  - , Stephen Brett
  - , Anthony Gordon
  - , Maie Templeton
  - , David Antcliffe
  - , Dorota Banach
  - , Sarah Darnell
  - , Ziortza Fernandez

    - , Eleanor Jepson
    - , Amal Mohammed
    - , Roceld Rojo
    - , Sonia Sousa Arias
    - , Anita Tamang Gurung
    - & Jenny Wong

- **The Royal Liverpool University Hospital, Liverpool, UK**

    - Jaime Fernandez-Roman
    - , David O. Hamilton
    - , Emily Johnson
    - , Brian Johnston
    - , Maria Lopez Martinez
    - , Suleman Mulla
    - , Alicia A. C. Waite
    - , Karen Williams
    - , Victoria Waugh
    - , Ingeborg Welters
    - , Jessica Emblem
    - , Maria Norris
    - & David Shaw

- **John Radcliffe Hospital, Oxford, UK**

    - Archana Bashyal
    - , Sally Beer
    - , Paula Hutton
    - , Stuart McKechnie
    - , Neil Davidson
    - , Soya Mathew
    - , Grace Readion
    - , Jung Ryu
    - & Jean Wilson

- **Addenbrooke’s Hospital, Cambridge, UK**

- Shruti Agrawal
- , Kay Elston
- , Megan Jones
- , Eoghan Meaney
- , Petra Polgarova
- , Muhammad Elbehery
- , Charlotte Summers
- , Esther Daubney
- , Anthony Ng
- , Jocelyn Marshall
- , Nazima Pathan
- , Katerina Stroud
- & Deborah White

- **Nottingham University Hospital, Nottingham, UK**

  - Angela Andrew
  - , Saima Ashraf
  - , Amy Clark
  - , Martin Dent
  - , Margaret Langley
  - , Cecilia Peters
  - , Lucy Ryan
  - , Julia Sampson
  - , Shuying Wei
  - , Alice Baddeley
  - , Megan Meredith
  - , Lucy Morris
  - , Alexandra Gibbons
  - & Lisa McLoughlin

- **St George's Hospital, London, UK**

  - Carlos Castro Delgado
  - , Victoria Clark
  - , Deborah Dawson

    - , Lijun Ding
    - , Georgia Durrant
    - , Obiageri Ezeobu
    - , Abiola Harrison
    - , William James Hurt
    - , Rebecca Kanu
    - , Ashley Kinch
    - , Susannah Leaver
    - , Ana Lisboa
    - , Jisha Mathew
    - , Kamal Patel
    - , Romina Pepermans Saluzzio
    - , John Rawlins
    - , Tinashe Samakomva
    - , Nirav Shah
    - , Christine Sicat
    - , Joana Texeira
    - , Joana Gomes De Queiroz
    - , Edna Fernandes Da Gloria
    - , Elena Maccacari
    - , Nikki Yun
    - , Soumendu Manna
    - , Sarah Farnell-Ward
    - , Maria Maizcordoba
    - , Maria Thanasi
    - & Hawakin Haji Ali

- **BHRUT (Barking Havering)—Queens Hospital and King George Hospital, Essex, UK**

    - Janice Hastings
    - , Lina Grauslyte
    - , Musarat Hussain
    - , Bobby Ruge
    - , Sam King
    - , Tatiana Pogreban
    - , Lace Rosaroso

- , Helen Smith
- , Mandeep-Kaur Phull
- , Nikkita Adams
- , George Franke
- , Aparna George
- , Erika Salciute
- , Joanna Wong
- , Karen Dunne
- , Luke Flower
- , Emma Sharland
- & Sukhmani Sra

- **Royal Infirmary of Edinburgh, Edinburgh, UK**
  - Gillian Andrew
  - , Marie Callaghan
  - , Lucy Barclay
  - , Lucy Marshall
  - , Kenneth Baillie
  - , Maria Amamio
  - , Sophie Birch
  - , Kate Briton
  - , Sarah Clark
  - , Katerine Doverman
  - , Dave Hope
  - , Corrienne Mcculloch
  - , Scott Simpson
  - & Jo Singleton

- **Kingston Hospital, Surrey, UK**
  - Rita Fernandez
  - , Meryem Allen
  - , David Baptista
  - , Rebecca Crowe
  - , Jonathan Fox

- , Jacyntha Khera
- , Adam Loveridge
- , India McKenley
- , Eriko Morino
- , Andres Naranjo
- , Denise O’Connor
- , Richard Simms
- , Kathryn Sollesta
- , Andrew Swain
- , Harish Venkatesh
- , Rosie Herdman-Grant
- & Anna Joseph

- **Queen Alexandra Hospital, Portsmouth, UK**
  - Angela Nown
  - , Steve Rose
  - , David Pogson
  - , Helen Boxall
  - , Lutece Brimfield
  - , Helen Claridge
  - , Zoe Daly
  - , Shenu George
  - & Andrew Gribbin

- **Royal Gwent Hospital, Newport, UK**
  - Yusuf Cheema
  - , Sean Cutler
  - , Owen Richards
  - , Anna Roynon-Reed
  - , Shiney Cherian
  - , Anne Emma Heron
  - , Gemma Williams
  - , Tamas Szakmany
  - , Abby Waters

- , Kim Collins
- , Jill Dunhill
- , Ffion Jones
- , Rebecca Morris
- , Lucy Ship
- & Amy Cardwell

- **Royal Blackburn Teaching Hospital, Blackburn, UK**
  - Syamlan Ali
  - , Ravi Bhatterjee
  - , Rachel Bolton
  - , Srikanth Chukkambotla
  - , Dabheoc Coleman
  - , Jack Dalziel
  - , Joseph Dykes
  - , Christopher Fine
  - , Bethan Gay
  - , Wendy Goddard
  - , Drew Goodchild
  - , Rhiannan Harling
  - , Muhammad Hijazi
  - , Sarah Keith
  - , Meherunnisa Khan
  - , Roseanna Matt
  - , Janet Ryan-Smith
  - , Samuel Saad
  - , Philippa Springle
  - , Jacqueline Thomas
  - , Nick Truman
  - , Aayesha Kazi
  - , Matthew Smith
  - , Heather Collier
  - , Chloe Davison
  - , Stephen Duberley
  - , Jeanette Hargreaves
  - , Janice Hartley

    - , Tahera Patel
    - & Ellen Smith

- **Stepping Hill Hospital, Stockport, UK**

    - Alissa Kent
    - , Emma Goodwin
    - , Ahmed Zaki
    - , Clare Tibke
    - , Susan Hopkins
    - , Hywel Gerrard
    - , Matthew Jackson
    - , Sara Bennett
    - , Liane Marsh
    - & Rebecca Mills

- **Northumbria Healthcare NHS Foundation Trust, North Shields, UK**

    - Jessica Bell
    - , Helen Campbell
    - , Angela Dawson
    - , Steve Dodds
    - , Stacey Duffy
    - , Lisa Gallagher
    - , Gemma McCafferty
    - , Stacey Short
    - , Tracy Smith
    - , Kirsty Thomas
    - , Claire Walker
    - , Jessica Reynolds
    - , Bryan Yates
    - , Hayley McKie
    - , Maria Panteli
    - , Maria Thompson
    - & Gail Waddell

- **Countess of Chester Hospital, Chester, UK**

    - Sarah De Beger
    - , Azmerelda Abraheem
    - , Charlie Dunmore
    - , Rumanah Girach
    - , Rhianna Jones
    - , Emily London
    - , Imrun Nagra
    - , Farah Nasir
    - , Hannah Sainsbury
    - , Clare Smedley
    - , Stephen Brearey
    - , Caroline Burchett
    - , Kathryn Cawley
    - , Maria Faulkner
    - , Helen Jeffrey
    - , Peter Bamford
    - , Firdaus Shaikh
    - , Lauren Slack
    - & Angela Davies

- **Pinderfields General Hospital, Wakefield, UK**

    - Hollie Brooke
    - , Jose Cebrian Suarez
    - , Ruth Charlesworth
    - , Karen Hansson
    - , John Norris
    - , Alice Poole
    - , Rajdeep Sandhu
    - , Elizabeth Smithson
    - , Muthu Thirumaran
    - , Veronica Wagstaff
    - , Sarah Buckley
    - , Brendan Sloan
    - , Alastair Rose

- , Amy Major
- & Alexandra Metcalfe

- **Ninewells Hospital, Dundee, UK**

  - Christine Almaden-Boyle
  - , Pauline Austin
  - , Susan Chapman
  - , Alexandre Eros
  - , Louise Cabrelli
  - , Stephen Cole
  - , Clare Whyte
  - & Matt Casey

- **Croydon University Hospital, Croydon, UK**

  - Vasileios Bafitis
  - , George Tsinaslanidis
  - , Cassandra George
  - , Reena Khade
  - , Christopher Black
  - & Sundar Raj Ashok

- **Morriston Hospital, Swansea, UK**

  - Sean Farley
  - , Elaine Brinkworth
  - , Rachel Harford
  - , Carl Murphy
  - , Marie Williams
  - , Luke Newey
  - , Hannah Toghill
  - , Sophie Lewis
  - , Tabitha Rees
  - , Ceri Battle
  - , Mark Baker

- , Jenny Travers
- & Karen Chesters

- **Queen Elizabeth University Hospital, Glasgow, UK**

  - Nicola Baxter
  - , Andrew Arnott
  - , Gordan McCreath
  - , Christopher McParland
  - , Laura Rooney
  - , Malcolm Sim
  - , Steven Henderson
  - , Lynn Abel
  - , Carol Dalton
  - , Sophie Kennedy-Hay
  - , Lynn O'Donohoe
  - , Megan O'Hare
  - , Izabela Orlikowska
  - & Natasha Parker

- **Broomfield Hospital, Chelmsford, UK**

  - Fiona McNeela
  - , Amanda Lyle
  - , Alistair Hughes
  - , Jayachandran Radhakrishnan
  - & Sian Gibson

- **Heartlands Hospital, Birmingham, UK**
    - Hollie Bancroft
    - , Mary Bellamy
    - , Jacqueline Daglish
    - , Salma Kadiri
    - , Faye Moore
    - , Joanne Rhodes
    - , Mirriam Sangombe
    - , Zhane Peterkin
    - , James Scriven
    - & Margaret Carmody
- **Royal Sussex County Hospital, Brighton, UK**
    - Juliet Cottle
    - , Emily Peasgood
    - , Laura Ortiz-Ruiz de Gordoa
    - , Claire Phillips
    - & Denise Skinner
- **York Hospital, York, UK**
    - Zoe Cinquina
    - , Kate Howard
    - , Rosie Joy
    - , Samantha Roche
    - , Isobel Birkinshaw
    - , Joseph Carter
    - , Jo Ingham
    - , Nicola Marshall
    - , Harriet Pearson
    - & Zoe Scott
- **Queen Elizabeth Hospital, Birmingham, UK**

- Jo Dasgin
- , Jaspret Gill
- , Annette Nilsson
- , Amy Bamford
- , Diana Hull
- , James Scriven
- , Nafeesah Ahmadhaider
- , Michelle Bates
- & Christopher McGhee

- **Royal Glamorgan Hospital, Pontyclun, UK**

  - Hannah Ellis
  - , Gwenllian Sera Howe
  - , Jayaprakash Singh
  - , Natalie Stroud
  - , Lisa Roche
  - , Ceri Lynch
  - , Bethan Deacon
  - , Carla Pothecary
  - , Justyna Smeaton
  - & Kevin Agravante

- **Barnet Hospital, London, UK**

  - Vinodh Krishnamurthy
  - , Cynthia Diaba
  - , Lincy John
  - , Lai Lim
  - & Rajeev Jha

- **Wythenshawe Hospital, Manchester, UK**

  - Jasmine Egan
  - , Timothy Felton
  - , Susannah Glasgow

- , Grace Padden
- , Ozerah Choudhr
- , Joanne Bradley-Potts
- , Stuart Moss
- , Saejohn Lingeswaran
- , Peter Alexander
- , Craig Brandwood
- , Sofia Fiouni
- , Luke Ward
- , Schvearn Allen
- , Jane Shaw
- & Christopher Smith

- **Medway Maritime Hospital, Gillingham, UK**
    - Oluronke Adanini
    - , Rebecca Collins
    - , Maines Msiska
    - , Linda Ofori
    - , Nikhil Bhatia
    - & Hayley Dolan

- **Royal Berkshire NHS Foundation Trust, Berkshire, UK**
    - Mark Brunton
    - , Jess Caterson
    - , Holly Coles
    - , Liza Keating
    - , Emma Tilney
    - , Nicola Jacques
    - , Matthew Frise
    - , Jennifer Armistead
    - , Shauna Bartley
    - , Parminder Bhuie
    - , Sabi Rai
    - & Gabriela Tomkova

- **Whiston Hospital, Prescot, UK**

    - Sandra Greer
    - , Karen Shuker
    - & Ascanio Tridente

- **The Royal Oldham Hospital, Manchester, UK**

    - Emma Dobson
    - , Jodie Hunt
    - , Redmond Tully
    - , Joy Dearden
    - , Andrew Drummond
    - , Prakash Kamath
    - , Emily Bullock
    - , Michelle Mulcahy
    - , Shelia Munt
    - , Grainne O’Connor
    - , Jennifer Philbin
    - , Chloe Rishton
    - , Chloe Scott
    - & Sarah Winnard

- **Chesterfield Royal Hospital Foundation Trust, Chesterfield, UK**

    - Nurkamalia Hasni
    - , Rachel Gascoyne
    - , Joanne Hawes
    - , Kelly Pritchard
    - , Lesley Stevenson
    - , Amanda Whileman
    - , Sarah Beavis
    - , Lauren Bishop
    - , Cindy Cart
    - , Katie Dale
    - , Mary Kelly-Baxter

- , Adam Mendelski
- , Emma Moakes
- , Rheanna Smith
- , Jan Woodward
- & Stephanie Wright

- **Aberdeen Royal Infirmary, Aberdeen, UK**
  - Angela Allan
  - , Adriana Botello
  - , Jade Liew
  - , Jasmine Medhora
  - , Erin Trumper
  - , Felicity Savage
  - , Teresa Scott
  - , Marc Place
  - & Callum Kaye

- **Royal Devon and Exeter Hospital, Exeter, UK**
  - Sarah Benyon
  - , Suzie Marriott
  - , Linda Park
  - , Helen Quinn
  - , Daisy Skyes
  - , Lily Zitter
  - , Kizzy Baines
  - , Elizabeth Gordon
  - , Samantha Keenan
  - & Andrew Pitt

- **Glasgow Royal Infirmary, Glasgow, UK**
  - Katharine Duffy
  - , Jane Ireland
  - , Gary Semple

- , Lynne Turner
- , Susanne Cathcart
- , Dominic Rimmer
- , Alex Puxty
- , Kathryn Puxty
- , Andrew Hurst
- , Jennifer Miller
- , Susan Speirs
- & Lauren Walker

- **Blackpool Victoria Hospital, Blackpool, UK**

    - Zena Bradshaw
    - , Joanna Brown
    - , Sarah Melling
    - , Stephen Preston
    - , Nicola Slawson
    - , Scott Warden
    - , Alanna Beasley
    - , Emma Stoddard
    - , Leonie Benham
    - , Jason Cupitt
    - , Melanie Caswell
    - , Lisa Elawamy
    - & Ashleigh Wignall

- **Southampton General Hospital, Southampton, UK**

    - Belinda Roberts
    - , Hannah Golding
    - , Samantha Leggett
    - , Michelle Male
    - , Martyna Marani
    - , Kirsty Prager
    - , Toran Williams
    - , Kim Golder

- , Oliver Jones
- , Rebecca Cusack
- , Clare Bolger
- , Rachel Burnish
- , Michael Carter
- , Susan Jackson
- , Karen Salmon
- & Jonathan Biss

- **Ashford and St Peter's Hospital, Surrey, UK**
  - Maia Aquino
  - , Maria Croft
  - , Victoria Frost
  - , Ian White
  - & Keshnie Govender

- **Derriford Hospital, Plymouth, UK**
  - Natasha Webb
  - , Liana Stapleton
  - , Colin Wells
  - , Nikitas Nikitas
  - , Ana Sanchez-Rodriguez
  - , Kayleigh Spencer
  - & Bethan Stowe

- **East Surrey Hospital, Redhill, UK**
  - Yvonne Izzard
  - , Michelle Poole
  - , Sonja Monnery
  - , Sallyanne Trotman
  - , Valerie Beech
  - , Edward Combes
  - & Teishel Joefield

- **Poole Hospital, Poole, UK**
    - Patrick Covernton
    - , Sarah Savage
    - , Elizabeth Woodward
    - , Julie Camsooksai
    - , Henrik Reschreiter
    - , Charlotte Barclay
    - , Yasmin DeAth
    - , Judith Dube
    - , Charlotte Humphrey
    - , Sarah Jenkins
    - , Emma Langridge
    - , Rebecca Milne
    - , Beverley Wadams
    - & Megan Woolcock
- **Royal Alexandra Hospital, Paisley, UK**
    - Michael Brett
    - , Brian Digby
    - , Lisa Gemmell
    - , James Hornsby
    - , Patrick MacGoey
    - , Pauline O'Neil
    - , Richard Price
    - , Radha Sundaram
    - , Lynn Abel
    - , Natalie Rodden
    - , Nicola Thomson
    - , Kevin Rooney
    - , Susan Currie
    - , Natasha Parker
    - , Lauren Walker
    - & Philip Henderson

- **St James’s University Hospital and Leeds General Infirmary, Leeds, UK**
  - Bethan Ogg
  - , Simon Whiteley
  - , Liz Wilby
  - , Kate Long
  - , Shailamma Matthew
  - , Sheila Salada
  - , Susan Trott
  - , Sarah Watts
  - , Zoe Friar
  - & Abigail Speight
- **Bedford Hospital, Bedford, UK**
  - Victoria Bastion
  - , Humza Chandna
  - , Brice Djeugam
  - , Muhammad Haseeb
  - , Harriet Kent
  - , Gamu Lubimbi
  - , Sophie Murdoch
  - , Alastair Thomas
  - , Beena David
  - , Rachel Lorusso
  - , Ana Vochin
  - , Melchizedek Penacerrada
  - & Retno Wulandari
- **Southport and Formby District General Hospital, Ormskirk, UK**
  - Charlotte Heath
  - , Srinivas Jakkula
  - , Anna Morris
  - , Ashar Ahmed

- , Arvind Nune
- , Claire Buttriss
- & Emma Whitaker

- **The Tunbridge Wells Hospital and Maidstone Hospital, Kent, UK**
  - Miriam Davey
  - , David Golden
  - , Amy Acklery
  - , Fabio Fernandes
  - , Bec Seaman
  - & Victoria Earl

- **Queen Elizabeth Hospital, Woolwich, London, UK**
  - Amy Collins
  - , Waqas Khaliq
  - , Rachel Adam
  - & Estefania Treus

- **North Manchester General Hospital, Manchester, UK**
  - Sarah Holland
  - , Jordan Alfonso
  - , Bethan Blackledge
  - , Michelle Bruce
  - , Laura Jayne Durrans
  - , Ayaa Eltayeb
  - , Jade Harris
  - , Samuel Hey
  - , Martin Hruska
  - , Thomas Lamb
  - , Joanne Rothwell
  - , Adele Fitzgerald
  - , Gabriella Lindergard
  - , Helen T-Michael

- , Tracey Duncan
- , Sharon Baxter-Dore
- , Lisa Cooper
- , Claire Fox
- , Jacinta Guerin
- , Tracey Hodgkiss
- & Karen Connolly

- **Royal Victoria Infirmary, Newcastle Upon Tyne, UK**

    - Paul McAlinden
    - , Victoria Bridgett
    - , Maggie Fearby
    - , A. Gulati
    - , Helen Hanson
    - , Sinead Kelly
    - , Louise McCormack
    - , Rachel Nixon
    - , Philip Robinson
    - , Victoria Slater
    - , Elaine Stephenson
    - , Andrea Webster
    - , K. Webster
    - , Carole Hays
    - , Anne Hudson
    - , Bijal Patel
    - , Ian Clement
    - , John Davis
    - , Sarah Francis
    - & Douglas Jerry

- **Hull Royal Infirmary, Hull, UK**

    - Caroline Abernathy
    - , Louise Foster
    - , Andrew Gratrix

    - , Llucia Cabral-Ortega
    - , Matthew Hines
    - , Victoria Martinson
    - , Elizabeth Stones
    - & Karen Winter

- **Manchester Royal Infirmary, Manchester, UK**

    - Esther Barrow
    - , Katharine Wylie
    - , Deborah Baines
    - , Katie Birchall
    - , Laurel Kolakaluri
    - , Richard Clark
    - , Anila Sukumaran
    - , Craig Brandwood
    - , Melanie Barker
    - , Deborah Paripoorani
    - , Lara Smith
    - & Charlotte Taylor

- **Royal Derby Hospital, Derby, UK**

    - Charlotte Downes
    - , Melanie Hayman
    - , Katie Riches
    - , Priya Daniel
    - , Deepak Subramanian
    - , Kathleen Holding
    - , Mary Hilton
    - , Carly McDonald
    - & Georgina Richardson

- **Aintree University Hospital, Liverpool, UK**

    - Georgia Halladay

    - , Peter Harding
    - , Amie Reddy
    - , Ian Turner-Bone
    - , Laura Wilding
    - , Robert Parker
    - , Michaela Lloyd
    - , Leanne Smith
    - & Charlie Kelly

- **Fairfield General Hospital, Bury, UK**

    - Maria Lazo
    - , Alan Neal
    - , Olivia Walton
    - , Julie Melville
    - , Jay Naisbitt
    - , Emily Bullock
    - & Rosane Joseph

- **Norfolk and Norwich University hospital (NNUH), Norwich, UK**

    - Sara Callam
    - , Lisa Hudig
    - , Jocelyn Keshet-Price
    - , Katie Stammers
    - , Karen Convery
    - , Georgina Randell
    - & Deirdre Fottrell-Gould

- **Milton Keynes University Hospital, Milton Keynes, UK**

    - Esther Mwaura
    - , Sara-Beth Sutherland
    - , Richard Stewart
    - , Louise Mew
    - & Lynn Wren

- **Good Hope Hospital, Birmingham, UK**

    - Laura Thrasyvoulou
    - , Heather Willis
    - , James Scriven
    - , Bridget Hopkins
    - , Daniel Lenton
    - & Abigail Roberts

- **Queen Elizabeth Hospital Gateshead, Gateshead, UK**

    - Maria Bokhari
    - , Rachael Lucas
    - , Wendy McCormick
    - , Jenny Ritzema
    - , Vanessa Linnett
    - , Amanda Sanderson
    - & Helen Wild

- **Royal Bolton Hospital, Bolton, UK**

    - Rebecca Flanagan
    - , Robert Hull
    - , Kat Rhead
    - , Emma McKenna
    - , Gareth Hughes
    - , Jennifer Anderson
    - , Kelly Jones
    - , Scott Latham
    - & Heather Riley

- **Tameside General Hospital, Ashton Under Lyne, UK**

    - Martina Coulding
    - , Martyn Clark
    - , Jacqueline McCormick

- , Oliver Mercer
    - , Darsh Potla
    - , Hafiz Rehman
    - , Heather Savill
    - , Victoria Turner
    - , Edward Jude
    - & Susan Kilroy

- **Salford Royal Hospital, Manchester, UK**

    - Elena Apetri
    - , Cathrine Basikolo
    - , Bethan Blackledge
    - , Laura Catlow
    - , Matthew Collis
    - , Reece Doonan
    - , Jade Harris
    - , Alice Harvey
    - , Karen Knowles
    - , Stephanie Lee
    - , Diane Lomas
    - , Chloe Lyons
    - , Liam McMorrow
    - , Angiy Michael
    - , Jessica Pendlebury
    - , Jane Perez
    - , Maria Poulaka
    - , Nicola Proudfoot
    - , Kathryn Slevin
    - , Vicky Thomas
    - , Danielle Walker
    - , Paul Dark
    - , Bethan Charles
    - , Danielle McLaughlan
    - , Melanie Slaughter
    - , Dan Horner
    - , Kathryn Cawley

- & Tracy Marsden

- **Great Ormond St Hospital and UCL Great Ormond St Institute of Child Health NIHR Biomedical Research Centre, London, UK**

  - Joyann Andrews
  - , Emily Beech
  - , Olugbenga Akinkugbe
  - , Alasdair Bamford
  - , Holly Belfield
  - , Gareth A. L. Jones
  - , Tara McHugh
  - , Hamza Meghari
  - , Samiran Ray
  - , Ana Luisa Tomas
  - , Lauran O'Neill
  - , Mark Peters
  - , Michael Bell
  - , Sarah Benkenstein
  - , Catherine Chisholm
  - , Charlene Davies
  - , Klaudia Kupiec
  - & Caroline Payne

- **Southmead Hospital, Bristol, UK**

  - Joanna Halls
  - , Hayley Blakemore
  - , Elizabeth Goff
  - , Kati Hayes
  - , Kerry Smith
  - , Deanna Stephens
  - , Ruth Worner
  - , Borislava Borislavova
  - , Beverley Faulkner
  - , Matt Thomas

- , Ruth Cookson
- , Emma Gendall
- , Georgina Larman
- , Rebecca Pope
- & Artur Smalira

- **William Harvey Hospital, Ashford, UK**

  - Victoria Priestley
  - , Tracey Cosier
  - , Gemma Millen
  - , James Rand
  - , Natasha Schumacher
  - , Roxana Sandhar
  - , Heather Weston
  - , Neil Richardson
  - & Lucy Cooper

- **Arrowe Park Hospital, Wirral, UK**

  - Cathy Jones
  - , Ya-Wen Jessica Huang
  - , Reni Jacob
  - , Craig Denmade
  - & Lewis McIntyre

- **Royal Hampshire County Hospital, Hampshire, UK**

  - Dawn Trodd
  - , Jane Martin
  - , Geoff Watson
  - , Emily Bevan
  - & Caroline Wreybrown

- **Bradford Royal Infirmary, Bradford, UK**

- Shereen Bano
- , Ruth Bellwood
- , Michael Bentley
- , Matt Bromley
- , Lucy Gurr
- , Camilla Ledgard
- , Janet McGowan
- , Kate Pye
- , Kirsten Sellick
- , Amelia Stacey
- , Deborah Warren
- , Brian Wilkinson
- , Louise Akeroyd
- , Huma Shafique
- , James Morgan
- , Susan Shorter
- , Rachel Swinger
- , Emily Waters
- & Tom Lawton

- **Glan Clwyd Hospital, Bodelwyddan, UK**

  - Elizabeth Allan
  - , Kate Darlington
  - , Ffyon Davies
  - , Llinos Davies
  - , Jack Easton
  - , Sumit Kumar
  - , Richard Lean
  - , Callum Mackay
  - , Richard Pugh
  - , Xinyi Qiu
  - , Stephanie Rees
  - , Jeremy Scanlon
  - , Joanne Lewis
  - , Daniel Menzies
  - , Annette Bolger

    - , Gwyneth Davies
    - , Jennifer Davies
    - , Esther Garrod
    - , Helen Jones
    - , Rachel Manley
    - & Hannah Williams

- **Royal Bournemouth Hospital, Bournemouth, UK**

    - Jordan Frankham
    - , Sally Pitts
    - , Nigel White
    - , Debbie Branney
    - & Heather Tiller

- **Bristol Royal Infirmary, Bristol, UK**

    - Georgia Efford
    - , Zoe Garland
    - , Lisa Grimmer
    - , Bethany Gumbrill
    - , Rebekah Johnson
    - , Katie Sweet
    - , Jeremy Bewley
    - , Christina Coleman
    - , Katie Corcoran
    - , Eva Maria Hernandez Morano
    - , Rachel Shiel
    - , Denise Webster
    - , Josephine Bonnici
    - , Eleanor Daniel
    - & Abbie Dell

- **University Hospital North Durham, Darlington, UK and Darlington Memorial Hospital, Darlington, UK**

- Melanie Kent
- , Ami Wilkinson
- , Ellen Brown
- , Andrea Kay
- , Suzanne Campbell
- , Amanda Cowton
- , Mark Birt
- , Vicki Greenaway
- , Kathryn Potts
- , Clare Hutton
- & Andrew Shepperson

- **Basildon Hospital, Basildon, UK**

  - Miranda Forsey
  - , Alice Nicholson
  - , Mark Vertue
  - , Joanne Riches
  - , Agilan Kaliappan
  - & Anne Nicholson

- **University College Hospital, London, UK**

  - Niall MacCallum
  - , Eamon Raith
  - , Georgia Bercades
  - , Ingrid Hass
  - , David Brealey
  - , Gladys Martir
  - , Anna Reyes
  - , Deborah Smyth
  - & Maria Zapatamartinez

- **Whittington Hospital, London, UK**

  - Ana Alvaro

    - , Champa Jetha
    - , Louise Ma
    - , Lauren Booker
    - , Loreta Mostoles
    - , Anezka Pratley
    - , Abdelhakim Altabaibeh
    - , Chetan Parmar
    - & Kayleigh Gilbert

- **Western General Hospital, Edinburgh, UK**

    - Susie Ferguson
    - , Amy Shepherd
    - , Sheila Morris
    - , Jo Singleton
    - , Rosie Baruah
    - , Maria Amamio
    - , Sophie Birch
    - , Kate Briton
    - , Sarah Clark
    - , Katherine Doverman
    - , Lucy Marshall
    - & Scott Simpson

- **Ipswich Hospital, Ipswich, UK**

    - Georgina Lloyd
    - , Stephanie Bell
    - , Vanessa Rivers
    - & Bally Purewal

- **Hereford County Hospital, Hereford, UK**

    - Kate Hammerton
    - , Susan Anderson
    - , Janine Birch

    - , Emma Collins
    - & Ryan Oleary

- **Sunderland Royal Hospital, Sunderland, UK**

    - Sarah Cornell
    - , Jordan Jarmain
    - , Kimberley Rogerson
    - , Fiona Wakinshaw
    - , Lindsey Woods
    - , Anthony Rostron
    - , Zeynep Elcioglu
    - & Alistair Roy

- **Queens Hospital Burton, Burton-On-Trent, UK**

    - Gillian Bell
    - , Holly Dickson
    - , Louise Wilcox
    - , Amro Katary
    - & Katy English

- **Musgrove Park Hospital, Taunton, UK**

    - Joanne Hutter
    - , Corinne Pawley
    - , Patricia Doble
    - , Charmaine Shovelton
    - , Marius Vaida
    - , Rebecca Purnell
    - & Ashly Thomas

- **The Royal Papworth Hospital, Cambridge, UK**

    - Lenka Cagova
    - , Adama Fofano

- , Helen Holcombe
- , Alice Michael Mitchell
- , Lucy Mwaura
- , Krithivasan P. Raman
- , Lucie Garnr
- , Sue Mepham
- , Kitty Paques
- , Alain Vuylsteke
- , Jennifer Mackie
- , Carmen Pearn
- & Julie Zamikula

- **University Hospital Lewisham, London, UK**
  - Mark Birt
  - , Estefania Treus Gude
  - , Maggie Nyirenda
  - , Lisa Capozzi
  - , Rosie Reece-Anthony
  - , Waqas Khaliq
  - , Hazma Noor
  - & Alfa Cresia Nilo

- **The Princess Alexandra Hospital, Harlow, UK**
  - Michelle Grove
  - , Amelia Daniel
  - , Amy Easthope
  - , Joanne Finn
  - , Nikki White
  - , Rajnish Saha
  - , Bibi Badal
  - & Karen Ixer

- **University Hospital of Wales, Cardiff, UK**

- Donna Duffin
- , Ben Player
- , Helen Hill
- , Jade Cole
- , Jenny Brooks
- , Michelle Davies
- , Rhys Davies
- , Lauren Hunt
- , Emma Thomas
- & Angharad Williams

- **West Middlesex Hospital, Isleworth, UK**

  - Metod Oblak
  - , Mini Thankachen
  - , Jamie Irisari
  - , Amrinder Sayan
  - & Monica Popescu

- **Royal Albert Edward Infirmary, Wigan, UK**

  - Cheryl Finch
  - , Andrew Jamieson
  - , Alison Quinn
  - , Joshua Cooper
  - , Sarah Liderth
  - & Natalia Waddington

- **Stoke Mandeville Hospital, Buckinghamshire, UK**

  - Iona Burn
  - , Katarina Manso
  - , Ruth Penn
  - , Julie Tebbutt
  - , Danielle Thornton
  - , James Winchester

    - , Geraldine Hambrook
    - & Pradeep Shanmugasundaram

- **Royal Lancaster Infirmary, Lancaster, UK**

    - Jayne Craig
    - , Kerry Simpson
    - , Andrew Higham
    - & Louise Sibbett

- **Basingstoke and North Hampshire Hospital, Basingstoke, UK**
    - Sheila Paine
    - , Annabel Reed
    - , Jo-Anna Conyngham
    - , McDonald Mupudzi
    - , Rachel Thomas
    - , Mary Wright
    - , Denise Griffin
    - , Richard Partridge
    - , Maria Alvarez Corral
    - , Nycola Muchenje
    - , Mildred Sitonik
    - & Caroline Wrey Brown
- **Worthing Hospital, Worthing, UK and St Richard's Hospital, Chichester, UK**
    - Aaron Butler
    - , Linda Folkes
    - , Heather Fox
    - , Amy Gardner
    - , David Helm
    - , Gillian Hobden
    - , Kirsten King
    - , Jordi Margalef
    - , Michael Margarson
    - , Tim Martindale
    - , Emma Meadows
    - , Dana Raynard
    - , Yvette Thirlwall
    - , Yolanda Baird
    - , Raquel Gomez
    - , Darren Martin
    - , Luke Hodgson

- , Clinton Corin
- , Erikka Sidall
- , Densie Szabo
- & Sharon Floyd

- **The Alexandra Hospital, Redditch and Worcester Royal Hospital, Worcester, UK**

  - Hannah Davies
  - , Karen Austin
  - , Olivia Kelsall
  - , Hannah Wood
  - , Hannah Davies
  - , Peter Anderson
  - , Katie Archer
  - , Andrew Burtenshaw
  - , Sarah Clayton
  - , Naiara Cother
  - , Nicholas Cowley
  - , Caroline Davis
  - , Stephen Digby
  - , Alison Durie
  - , Alison Harrison
  - , Emma Low
  - , Michael McAlindon
  - , Alex McCurdy
  - , Aled Morgan
  - , Tobias Rankin
  - , Jessica Thrush
  - , Helen Tranter
  - , Charlie Vigurs
  - , Laura Wild
  - , Karen Austin
  - , Olivia Kelsall
  - & Hannah Wood

- **Royal Cornwall Hospital, Truro, UK**

    - Thomas Cornell
    - , Kate Ralph
    - , Sarah Bean
    - , Karen Burt
    - , Michael Spivey
    - , Carol Richards
    - & Rachel Tedstone

- **Watford General Hospital, Watford, UK**

    - Siobhain Carmody
    - , Xiaobei Zhao
    - , Valerie Page
    - , Mark Louie Guanco
    - , Elvira Hoxha
    - & Camilla Zorloni

- **Macclesfield District General Hospital, Macclesfield, UK**

    - Charlotte Dean
    - , Emma Jones
    - , Emma Carter
    - , Joshua Dunn
    - , Thomas Kong
    - , Mervin Mahenthran
    - , Chris Marsh
    - , Maureen Holland
    - , Natalie Keenan
    - , Mohamed Mahmoud
    - , Marc Lyons
    - , Joanne Bradley-Potts
    - , Helen Wassall
    - & Meghan Young

- **Royal Surrey County Hospital, Guildford, UK**

- Paul Bradley
- , Dorota Burda
- , Sinead Donlon
- , Lesley Harden
- , Celia Harris
- , Irving Mayangao
- , Rugia Montaser
- , Sheila Mtuwa
- , Charles Piercy
- , Eleanor Smith
- , Sarah Stone
- , Jerik Verula
- , Helen Blackman
- , Cheryl Marriott
- , Natalia Michalak
- , Ben Creagh-Brown
- , Armorel Salberg
- , Naomi Boyer
- & Veronika Pristopan

- **Rotherham General Hospital, Rotherham, UK**

  - Victoria Maynard
  - , Rachel Walker
  - , Anil Hormis
  - , Dawn Collier
  - , Cheryl Graham
  - , Vicky Maynard
  - , Jake McCormick
  - & Jake Warrington

- **Craigavon Area Hospital, County Armagh, Northern Ireland**

  - Denise Cosgrove
  - , Denise McFarland
  - , Judith Ratcliffe

- & Rob Charnock

- **King’s Mill Hospital, Nottingham, UK**
    - Inez Wynter
    - , Mandy Gill
    - , Jill Kirk
    - , Paul Paul
    - , Valli Ratnam
    - & Sarah Shelton

- **Dumfries and Galloway Royal Infirmary, Dumfries, UK**
    - Catherine Jardine
    - , Alasdair Hay
    - & Dewi Williams

- **Prince Charles Hospital, Merthyr Tydfil, UK**
    - Bethan Deacon
    - , Latha Durga
    - , Meg Hibbert
    - , Gareth Kennard-Holden
    - , Chrsitopher Woodford
    - , Carla Pothecary
    - , Lisa Roche
    - , Dariusz Tetla
    - , Kevin Agravante
    - & Justyna Smeaton

- **Ysbyty Gwynedd, Bangor, UK**
    - Alicia Price
    - , Alice Thomas
    - , Chris Thorpe
    - , Ellen Knights

    - & Donna Ward

- **Royal Preston Hospital, Preston, UK**

    - Shondipon Laha
    - , Mark Verlander
    - & Alexandra Williams

- **The Great Western Hospital, Swindon, UK**

    - Rachel Prout
    - , Helen Langton
    - , Malcolm Watters
    - , Charlotte Hunt
    - & Catherine Novis

- **Lincoln County Hospital, Lincoln, UK**

    - Sarwat Arif
    - , Amy Cunningham
    - , Claire Hewitt
    - , Julia Hindale
    - , Karen Jackson-Lawrence
    - , Sarah Shepardson
    - , Maryanne Wills
    - , Susie Butler
    - , Silivia Tavares
    - , Russell Barber
    - , Annette Hilldrith
    - & Kelly Hubbard

- **University Hospital of North Tees, Stockton on Tees, UK**

    - Dawn Egginton
    - , Michele Clark
    - , Sarah Purvis

- , Simon Sinclair
  - & Vicky Collins

- **Glangwili General Hospital, Camarthen, UK**

  - Bethan Landeg
  - , Craig Sell
  - , Samantha Coetzee
  - , Alistair Gales
  - , Igor Otahal
  - , Becky Icke
  - , Meena Raj
  - , Caroline Williams
  - , Jill Williams
  - & Lucy Hill

- **Southend University Hospital, Westcliff-on-Sea, UK**

  - Abdul Kayani
  - , Bridgett Masunda
  - , Prisca Gondo
  - & Nigara Atayeva

- **Lister Hospital, Stevenage, UK**

  - Carina Cruz
  - & Natalie Pattison

- **Diana Princess of Wales Hospital, Grimsby, UK**

  - Caroline Burnett
  - , Jonathan Hatton
  - , Elaine Heeney
  - , Maria Newton
  - , Hassan Al-Moasseb
  - , Teresa Behan

- , Jasmine Player
- , Rachael Stead
- , Atideb Mitra
- & Kirsty Nauyokas

- **West Suffolk Hospital, Bury St Edmunds, UK**
  - Sally Humphreys
  - , Helen Cockerill
  - & Ruth Tampsett

- **Victoria Hospital, Kirkcaldy, UK**
  - Evgeniya Postovalova
  - , Tina Coventry
  - , Amanda McGregor
  - , Susan Fowler
  - , Mike Macmahon
  - , Patricia Cochrane
  - & Sandra Pirie

- **Calderdale Royal Hospital, Halifax, UK and Huddersfield Royal Infirmary, Huddersfield, UK**
  - Sarah Hanley
  - , Asifa Ali
  - , Megan Brady
  - , Sam Dale
  - , Annalisa Dance
  - , Lisa Gledhill
  - , Jill Greig
  - , Kathryn Hanson
  - , Kelly Holdroyd
  - , Marie Home
  - , Tahira Ishaq
  - , Diane Kelly

- , Lear Matapure
- , Deborah Melia
- , Samantha Mellor
- , Ekta Merwaha
- , Tonicha Nortcliffe
- , Lisa Shaw
- , Ryan Shaw
- , Tracy Wood
- , Lee-Ann Bayo
- , Miranda Usher
- , Alison Wilson
- , Ross Kitson
- , Jez Pinnell
- , Matthew Robinson
- & Kaitlin Boltwood

- **Dorset County Hospital, Dorchester, UK**

  - Jenny Birch
  - , Laura Bough
  - , Rebecca Tutton
  - , Barbara Winter-Goodwin
  - , Josie Goodsell
  - , Kate Taylor
  - , Patricia Williams
  - , Sarah Williams
  - , Ashleigh Cave
  - & James Rees

- **Russell's Hall Hospital, Dudley, UK**

  - Janet Imeson-Wood
  - , Jacqueline Smith
  - , Vishal Amin
  - , Komala Karthik
  - , Rizwana Kausar

    - , Elena Anastasescu
    - , Karen Reid
    - , Vikram Anumakonda
    - & Ella Stoddart

- **Royal United Hospital, Bath, UK**

    - Carrie Demetriou
    - , Charlotte Eckbad
    - , Lucy Howie
    - , Sarah Mitchard
    - , Lidia Ramos
    - , Katie White
    - , Sarah Hierons
    - , Fiona Kelly
    - , Alfredo Serrano-Ruiz
    - & Gabrielle Evans

- **St Mary's Hospital, Newport, UK**

    - Liz Nicol
    - , Joy Wilkins
    - , Kim Hulacka
    - , Gabor Debreceni
    - , Alison Brown
    - & Vikki Crickmore

- **George Eliot Hospital NHS Trust, Nuneaton, UK**

    - Kay Hill
    - & Thogulava Kannan

- **Yeovil Hospital, Yeovil, UK**

    - Zenaida Dagutao
    - , Kate Beesley

    - , Alison Lewis
    - , Jess Perry
    - , Sherly Antony
    - , Sarah Board
    - , Clare Buckley
    - , Lucy Pippard
    - , Alfonso Tanate
    - , Diane Wood
    - , Agnieska Kubisz-Pudelko
    - & Ayman Gouda

- **Forth Valley Royal Hospital, Falkirk, UK**

    - Fiona Auld
    - , Joanne Donnachie
    - , Euan Murdoch
    - , Lynn Prentice
    - , Nikole Runciman
    - , Dhaneesha Senaratne
    - , Abigail Short
    - , Laura Sweeney
    - , Lesley Symon
    - , Anne Todd
    - , Patricia Turner
    - , Erin McCann
    - , Dario Salutous
    - , Ian Edmond
    - & Lesley Whitelaw

- **Frimley Park Hospital, Surrey, UK**

    - Harish Venkatesh
    - , Yvonne Bland
    - , Istvan Kajtor
    - , Lisa Kavanagh
    - , Karen Singler

- & George Linfield-Brown

- **Chelsea & Westminster NHS Foundation Trust, London, UK**
  - Luke Stephen Prockter Moore
  - , Marcela Vizcaychipi
  - , Laura Martins
  - , Luke Moore
  - , Rhian Bull
  - & Jaime Carungcong

- **Queen Elizabeth the Queen Mother Hospital, Margate, UK**
  - Louise Allen
  - , Eva Beranova
  - , Alicia Knight
  - , Carly Price
  - , Sorrell Tilbey
  - , Sharon Turney
  - , Tracy Hazelton
  - , Gabriella Tutt
  - , Mansi Arora
  - , Salah Turki
  - , Emily Sinfield
  - , Joanne Deery
  - & Hazel Ramos

- **Royal Brompton Hospital, London, UK**
  - Daniele Cristiano
  - , Natalie Dormand
  - , Zohreh Farzad
  - , Mahitha Gummadi
  - , Sara Salmi
  - , Geraldine Sloane
  - , Mathew Varghese

    - , Vicky Thwaites
    - , Brijesh Patel
    - , Liyanage Kamal
    - & Anelise Catelan Zborowski

- **Darent Valley Hospital, Dartford, UK**

    - Ryan Coe
    - , Madeleine Anderson
    - , Jane Beadle
    - , Charlotte Coates
    - , Katy Collins
    - , Maria Crowley
    - , Laura Johnson
    - , Laura King
    - , Remi Paramsothy
    - , Janet Sargeant
    - , Pedro Silva
    - , Carmel Stuart
    - , June Taylor
    - , David Tyl
    - , Phillipa Wakefield
    - , Charlotte Kamundi
    - , Olumide Olufuwa
    - , Zakaulla Belagodu
    - , Anca Gherman
    - & Naomi Oakley

- **University Hospital Crosshouse, Kilmarnock, UK**

    - John Allan
    - , Tim Geary
    - , Alistair Meikle
    - , Peter O'Brien
    - , Stephen Wood
    - , Andrew Clark

- & Gordon Houston

- **University Hospital Wishaw, Wishaw, UK**
  - Karen Black
  - , Michelle Clarkson
  - , Stuart D’Sylva
  - , Alan Morrison
  - , Kathryn Norman
  - , Margaret Taylor
  - , Suzanne Clements
  - , Catriona Cohrane
  - , Nora Gonzalez
  - , Dominic Strachan
  - , Claire Beith
  - & Kirsten Moar

- **University College Dublin, St Vincent’s University Hospital, Dublin, Ireland**
  - Lorna Murphy
  - , Michelle Smythe
  - , Alistair Nichol
  - & Kathy Brickell

- **The Queen Elizabeth Hospital, King’s Lynn, UK**
  - Inthakab Ali Mohamed Ali
  - , Karen Beaumont
  - , Mohamed Elsaadany
  - , Kay Fernandes
  - , Sameena Mohamed Ally
  - , Harini Rangarajan
  - , Varun Sarathy
  - , Sivarupan Selvanayagam
  - , Dave Vedage

    - , Matthew White
    - , Zoe Coton
    - , Aricsa Joshy
    - , Mark Blunt
    - & Hollie Curgenven

- **Walsall Manor Hospital, Walsall, UK**

    - Liam Botfield
    - , Catherine Dexter
    - , Aditya Kuravi
    - , Joanne Butler
    - , Robert Chadwick
    - , Poonam Ranga
    - , Lisa Richardson
    - , Emma Virgilio
    - , Maddiha Anwer
    - , Atul Garg
    - , Donna Botfield
    - & Xana Marriott

- **Princess Royal Hospital, Brighton, UK**

    - Keely Stewart
    - , Dee Mullan
    - , Claire Phillips
    - , Jane Gaylard
    - , Justyna Nowak
    - & Denise Skinner

- **Barnsley Hospital, Barnsley, UK**

    - Sian Jones
    - , Rikki Crawley
    - , Abigail Crew
    - , Mishell Cunningham

- , Allison Daniels
- , Laura Harrison
- , Susan Hope
- , Nicola Lancaster
- , Jamie Matthews
- , Gemma Wray
- , Alice Nicholson
- , Ken Inweregbu
- , Sarah Cutts
- & Katharine Miller

- **Warrington General Hospital, Warrington, UK**

  - Ailbhe Brady
  - , Rebekah Chan
  - , Shane McIvor
  - , Helena Prady
  - , Bijoy Mathew
  - , Jeff Little
  - & Tim Furniss

- **Royal Victoria Hospital, Belfast, NI**

  - Chris Wright
  - , Bernadette King
  - , Christopher Wasson
  - , Aisling O'Neill
  - , Christine Turley
  - , Peter McGuigan
  - , Erin Collins
  - , Stephanie Finn
  - , Jackie Green
  - , Julie McAuley
  - , Abitha Nair
  - , Charlotte Quinn
  - , Suzanne Tauro

    - , Kathryn Ward
    - , Michael McGinlay
    - & Kiran Reddy

- **Royal Hallamshire Hospital and Northern General Hospital, Sheffield, UK**

    - Norfaizan Ahmad
    - , Samantha Anderson
    - , Joann Barker
    - , Kris Bauchmuller
    - , Kathryn Birchall
    - , Sarah Bird
    - , Kay Cawthron
    - , Luke Chetam
    - , Joby Cole
    - , Ben Donne
    - , David Foote
    - , Amber Ford
    - , Helena Hanratty
    - , Kate Harrington
    - , Lisa Hesseldon
    - , Kay Housley
    - , Yvonne Jackson
    - , Claire Jarman
    - , Faith Kibutu
    - , Becky Lenagh
    - , Irene Macharia
    - , Shamiso Masuko
    - , Leanne Milner
    - , Helen Newell
    - , Lorenza Nwafor
    - , Simon Oxspring
    - , Patrick Phillips
    - , Ajay Raithatha
    - , Sarah Rowland-Jones
    - , Jacqui Smith

- , Roger Thompson
- , Helen Trower
- , Sara Walker
- , James Watson
- , Matthew Wiles
- , Alison Lye
- , Jayne Willson
- , Gary Mills
- , Sansha Harris
- & Eleanor Hartill

- **Harefield Hospital, London, UK**

  - Anthony Barron
  - , Ciara Collins
  - , Sundeep Kaul
  - , Claire Nolan
  - , Oliver Polgar
  - , Claire Prendergast
  - , Paula Rogers
  - , Rajvinder Shokkar
  - , Meriel Woodruff
  - , Kanta Mahay
  - , Vicky Thwaites
  - , Anna Reed
  - , Hayley Meyrick
  - , Heather Passmore
  - & James Farwell

- **Cumberland Infirmary, Carlisle, UK**

  - Alison Brown
  - , Susan O’Connell
  - , Jane Gregory
  - , Luigi Barberis
  - , Rosemary Harper

    - , Tim Smith
    - & Diane Armstrong

- **Eastbourne District General Hospital, East Sussex, UK and Conquest Hospital, East Sussex, UK**

    - Angie Bowey
    - , Anne Cowley
    - , Andrew Corner
    - , Judith Highgate
    - , Claire Rutherfurd
    - , Jo-Anne Taylor
    - , Sarah Goodwin
    - & Claire Rutherford

- **Salisbury District Hospital, Salisbury, UK**

    - Beena Eapen
    - , Fiona Trim
    - & Phil Donnison

- **Airedale General Hospital, Keighley, UK**

    - Lisa Armstrong
    - , Hayley Bates
    - , Emma Dooks
    - , Fiona Farquhar
    - , Amy Kitching
    - , Chantal McParland
    - , Sophie Packham
    - & Brigid Hairsine

- **Leicester Royal Infirmary, Leicester, UK**

    - Premetie Andreou
    - , Dawn Hales

- , Megha Mathews
- , Rekha Patel
- , Peter Barry
- , Neil Flint
- , Jessica Hailstone
- , Navneet Ghuman
- , Bethany Leonard
- & Rachel Lees

- **Peterborough City Hospital, Peterborough, UK and Hinchingbrooke Hospital, Huntingdon, UK**

  - Deborah Butcher
  - , Katy Leng
  - , Nicola Butterworth-Cowin
  - & Susie O'Sullivan

- **Colchester General Hospital, Colchester, UK**

  - Alison Ghosh
  - & Emma Williams

- **Princess Royal Hospital, Telford and Royal Shrewsbury Hospital, Shrewsbury, UK**

  - Colene Adams
  - , Anita Agasou
  - , Tracie Arden
  - , Mandy Beekes
  - , Amy Bowes
  - , Pauline Boyle
  - , Heather Button
  - , Mandy Carnahan
  - , Anne Carter
  - , Danielle Childs
  - , Jane Gaylard

    - , Fran Hurford
    - , Yasmin Hussain
    - , Ayesha Javaid
    - , James Jones
    - , Michael Leigh
    - , Terry Martin
    - , Helen Millward
    - , Nichola Motherwell
    - , Dee Mullan
    - , Julie Newman
    - , Rachel Rikunenko
    - , Jo Stickley
    - , Julie Summers
    - , Louise Ting
    - , Helen Tivenan
    - , Denise Donaldson
    - , Nigel Capps
    - , Emily Cale
    - , Sanal Jose
    - , Wendy Osbourne
    - , Susie Pajak
    - , Jayne Rankin
    - & Louise Tonks

- **University Hospital Monklands, Airdrie, UK**

    - Tracy Baird
    - , Margaret Harkins
    - , Jim Ruddy
    - & Joe West

- **Wrexham Maelor Hospital, Wrexham, Wales**

    - Joseph Duffield
    - , Lewis Mallon
    - , Oliver Smith

- , Sara Smuts
- , Andy Campbell
- , Cate Davies
- , Sarah Davies
- , Rachel Hughes
- , Lisa Jobes
- , Victoria Whitehead
- & Clare Watkins

- **Royal Hospital for Children, Glasgow, UK**

    - Fiona Bowman
    - , Barry Milligan
    - , Colin Begg
    - & Liane McPherson

- **New Cross Hospital, Wolverhampton, UK**

    - Stella Metherell
    - , Nichola Harris
    - , Victoria Lake
    - , Elizabeth Radford
    - , Andy Smallwood
    - , Shameer Gopal
    - & Katherine Vassell

- **University Hospital Hairmyres, East Kilbride, UK**

    - Dina Bell
    - , Rosalind Boyle
    - , Katie Douglas
    - , Lynn Glass
    - , Liz Lennon
    - , Austin Rattray
    - , Claire Beith
    - & Emma Lee

- **Warwick Hospital, Warwick, UK**
    - Danielle Jones
    - , Penny Parsons
    - , Ben Attwood
    - , Paul Jefferson
    - , Mohan Ranganathan
    - , Inderjit Atwal
    - , Bridget Campbell
    - , Angela Day
    - & Camilla Stagg
- **Sandwell General Hospital and City Hospital, Birmingham, UK**
    - Emma Haynes
    - , Cecilia Ahmed
    - , Sarah Clamp
    - , Julie Colley
    - , Risna Haq
    - , Anne Hayes
    - , Sibet Joseph
    - , Zahira Maqsood
    - , Samia Hussain
    - , Jonathan Hulme
    - , Patience Domingos
    - , Rita Kumar
    - , Manjit Purewal
    - & Becky Taylor
- **Royal Manchester Children’s Hospital, Manchester, UK**
    - Lara Bunni
    - , Monica Latif
    - , Claire Jennings
    - , Shilu Jose
    - , Rebecca Marshall

    - , Aleksandra Metryka
    - & Gayathri Subramanian

- **Gloucestershire Royal Hospital, Gloucester, UK**

    - Adam Burgoyne
    - , Susan O’Connell
    - , Amanda Tyler
    - , Joanne Waldron
    - , Paula Hilltout
    - & Jayne Evitts

- **University Hospitals Coventry & Warwickshire NHS Trust, Coventry, UK**

    - Geraldine Ward
    - , Pamela Bremmer
    - , Carl Hawkins
    - , Sophie Jackman
    - & Michal Ogorek

- **Torbay Hospital, Torquay, UK**

    - Kylie Ashby
    - , Lorraine Thornton
    - , Pauline Mercer
    - , Matthew Halkes
    - & Adam Revill

- **Pilgrim Hospital, Lincoln, UK**

    - Bryony Saint
    - , Jo Fletcher
    - , Kimberley Netherton
    - , Manish Chablani
    - , Amy Kirkby

- , Amanda Roper
- & Kinga Szymiczek

- **Prince Philip Hospital, Lianelli, UK**
  - Isobel Sutherland
  - , Linda O'Brien
  - , Igor Otahal
  - , Joanne Connell
  - , Kim Davies
  - , Tracy Lewis
  - , Zohra Omar
  - & Emma Perkins

- **Princess of Wales Hospital, Llantrisant, UK**
  - Lisa Roche
  - , Sonia Sathe
  - & Ellie Davies

- **Northampton General Hospital NHS Trust, Northampton, UK**
  - Alex Lyon
  - , Isheunesu Mapfunde
  - , Charlotte Willis
  - , Rachael Hitchcock
  - , Kathryn Hall
  - & Christopher King

- **The Christie NHS Foundation Trust, Manchester, UK**
  - Andrew Fagan
  - , Roonak Nazari
  - , Lucy Worsley
  - , Suzanne Allibone
  - , Vidya Kasipandian

- , Amit Patel
- , Parisa Cutting
- , Roman Genetu
- , Ainhi Mac
- , Anthony Murphy
- , Sinead Ward
- & Fatima Butt

- **James Paget University Hospital NHS Trust, Great Yarmouth, UK**

  - Amanda Ayers
  - , Wendy Harrison
  - , Katherine Mackintosh
  - & Julie North

- **Birmingham Children's Hospital, Birmingham, UK**

  - Lydia Ashton
  - , Rehana Bi
  - , Samantha Owen
  - , Helen Winmill
  - & Barney Scholefield

- **Withybush General Hospital, Pembrokeshire, Wales**

  - Hannah Blowing
  - , Erin Williams
  - , Michaela Duskova
  - , Michelle Edwards
  - , Alun Rees
  - , Helen Thomas
  - , Rachel Hughes
  - , Igor Otahal
  - , Jolene Brooks
  - , Janet Phipps
  - & Suzanne Brooks

- **Northwick Park Hospital, London, UK**

  - Catherine Dennis
  - , Vicki Parris
  - , Sinduya Srikaran
  - , Anisha Sukha
  - , Alistair McGregor
  - & Gerlynn Tiongson

- **North Devon District Hospital, Barnstaple, UK**

  - Katie Adams
  - , Benedict Andrew
  - , Adam Brayne
  - , Sasha Carter
  - , Louise Findlay
  - , Emma Fisher
  - , Peter Jackson
  - , Duncan Kaye
  - , Juliet Parkin
  - , Victoria Tuckey
  - , Jane Hunt
  - , Nicholas Love
  - , Lynne van Koutrick
  - & Ashley Hanson

- **Scunthorpe General Hospital, Scunthorpe, UK**

  - Kathy Dent
  - , Elizabeth Horsley
  - , Sandra Pearson
  - , Sue Spencer
  - , Dorothy Hutchinson
  - , Jasmine Player
  - , Dorota Potoczna
  - , Muhammad Nauman Akhtar

    - , Lisa-Jayne Cottam
    - , Kirsty Nauyokas
    - & Jack Sanders

- **Royal Free Hospital, London, UK**

    - Sara Mingo Garcia
    - , Glykeria Pakou
    - , Cynthia Diaba
    - , Helder Filipe
    - , Lincy John
    - , Amitaa Maharajh
    - , Mark de Neef
    - , Daniel Martin
    - , Christine Eastgate
    - & Poh Choo Teoh

- **Raigmore Hospital, Inverness, UK**

    - Fiona Barrett
    - , Clare Bradley
    - , Avril Donaldson
    - , Mairi Mascarenhas
    - , Marianne O'Hara
    - , Laura Okeefe
    - , Noreen Clarke
    - , Jonathan Whiteside
    - , Rachael Campbell
    - , Joanna Matheson
    - , Deborah McDonald
    - & Donna Patience

- **West Cumberland Hospital, Whitehaven, UK**

    - Polly Rice
    - , Tim Smith

    - , Melanie Clapham
    - , Rachel Mutch
    - , Luigi Barberis
    - , Rosemary Harper
    - , Hannah Craig
    - & Una Poultney

- **Furness General Hospital, Barrow-in-Furness, UK**

    - Karen Burns
    - & Andrew Higham

- **Liverpool Heart and Chest Hospital, Liverpool, UK**

    - Sophie Twiss
    - , Janet Barton
    - , Linsha George
    - , Clare Harrop
    - , Sherly Mathew
    - & David Justin Wright

- **Scarborough General Hospital, Scarborough, UK**

    - Rachel Harrison
    - , Jordan Toohie
    - , Ben Chandler
    - , Alison Turnbull
    - , Janine Mallinson
    - & Kerry Elliott

- **Bronglais General Hospital, Aberystwyth, UK**

    - Rebecca Wolf-Roberts
    - , Helen Tench
    - , Igor Otahal
    - , Maria Hobrok

- , Ronda Loosley
- , Heather McGuinness
- & Tanya Sims

- **Alder Hey Children's Hospital, Liverpool, UK**

  - Deborah Afolabi
  - , Kathryn Sian Allison
  - , Taya Anderson
  - , Rachael Dore
  - , Dawn Jones
  - , Naomi Rogers
  - , Paula Saunderson
  - , Jennifer Whitbread
  - , Laura O'Malley
  - , Laura Rad
  - & Daniel Hawcutt

- **Borders General Hospital, Melrose, UK**

  - Jonathan Aldridge
  - , Melanie Tolson
  - & Sweyn Garrioch

- **Leighton Hospital, Cheshire, UK**

  - Joanne Tomlinson
  - & Michael Grosdenier

- **Kent & Canterbury Hospital, Canterbury, UK**

  - David Loader
  - , Ritoo Kapoor
  - & Gemma Hector

- **Harrogate and District NHS Foundation Trust, Harrogate, UK**

- Joslan Scherewode
- , Chunda Sri-Chandana
- , Lorraine Stephenson
- & Sarah Marsh

- **The Royal Marsden Hospital, London, UK**

    - Arnold Dela Rosa
    - , Shaman Jhanji
    - , Thomas Bemand
    - , Ryan Howle
    - , Ravishankar Rao Baikady
    - , Benjamin Thomas
    - , Ethel Black
    - & Kate Tatham

- **Ealing Hospital, Southall, UK**

    - Sambasivarao Gurram
    - , Ekaterina Watson
    - , Vicki Parris
    - , Sheena Quaid
    - & Alistair McGregor

- **St John's Hospital Livingston, Livingston, UK**

    - Anne Saunderson
    - , Rachel O'Brien
    - , Sam Moultrie
    - , Jen Service
    - , Clare Cheyne
    - , Miranda Odam
    - & Alison Wiliams

- **Wexham Park Hospital, Slough, UK**

- Nicky Barnes
- , Peter Csabi
- , Joana Da Rocha
- & Louika Glynou

- **Sheffield Children's Hospital, Sheffield, UK**

  - Amy Huffenberger
  - , Jade Bryant
  - , Amy Pickard
  - , Nicholas Roe
  - , Arianna Bellini
  - , Anton Mayer
  - , Amy Burrow
  - , Natalie Colley
  - , Jayne Evans
  - , Alex Howlett
  - & Zeinab Khalifeh

- **Homerton University Hospital Foundation NHS Trust, London UK**

  - Jerldine Pryce
  - , Claire Gorman
  - , Amy Easthope
  - , Rebecca Brady
  - , Elizabeth Timlick
  - , Pierre Antoine
  - & Abhinhav Gupta

- **National Hospital for Neurology and Neurosurgery, London, UK**

  - John Hardy
  - , Henry Houlden
  - , Eleanor Moncur
  - , Arianna Tucci

- , Eamon Raith
- , Ambreen Tariq
- & David Brealey

- **The Royal Alexandra Children's Hospital, Brighton, UK**

    - Emma Tagliavini
    - , Becky Ramsay
    - , Katy Fidler
    - , Kevin Donnelly
    - & Rebecca Hollis

- **Golden Jubilee National Hospital, Clydebank, UK**

    - Jocelyn Barr
    - , Elizabeth Boyd
    - , Val Irvine
    - , Ben Shelley
    - , Julie Buckley
    - , Charlene Hamilton
    - & Kathryn Valdeavella

SCOURGE Consortium

- Javier Abellan
- , René Acosta-Isaac
- , Jose María Aguado
- , Carlos Aguilar
- , Sergio Aguilera-Albesa
- , Abdolah Ahmadi Sabbagh
- , Jorge Alba
- , Sergiu Albu
- , Karla A. M. Alcalá-Gallardo
- , Julia Alcoba-Florez
- , Sergio Alcolea Batres
- , Holmes Rafael Algarin-Lara

- , Virginia Almadana
- , Julia Almeida
- , Berta Almoguera
- , María R. Alonso
- , Nuria Alvarez
- , Yady Álvarez-Benítez
- , Felipe Álvarez-Navia
- , Rodolfo Alvarez-Sala Walther
- , Álvaro Andreu-Bernabeu
- , Maria Rosa Antonijoan
- , Eunate Arana-Arri
- , Carlos Aranda
- , Celso Arango
- , Carolina Araque
- , Nathalia K. Araujo
- , Izabel M. T. Araujo
- , Ana C. Arcanjo
- , Ana Arnaiz
- , Francisco Arnalich Fernández
- , María J. Arranz
- , José Ramon Arribas Lopez
- , Maria-Jesus Artiga
- , Yubelly Avello-Malaver
- , Carmen Ayuso
- , Ana María Baldion
- , Belén Ballina Martín
- , Raúl C. Baptista-Rosas
- , Andrea Barranco-Díaz
- , María Barreda-Sánchez
- , Viviana Barrera-Penagos
- , Moncef Belhassen-Garcia
- , Enrique Bernal
- , David Bernal-Bello
- , Joao F. Bezerra
- , Marcos A. C. Bezerra
- , Natalia Blanca-López
- , Rafael Blancas

- , Lucía Boix-Palop
- , Alberto Borobia
- , Elsa Bravo
- , María Brion
- , Óscar Brochado-Kith
- , Ramón Brugada
- , Matilde Bustos
- , Alfonso Cabello
- , Juan J. Caceres-Agra
- , Esther Calbo
- , Enrique J. Calderón
- , Shirley Camacho
- , Marcela C. Campos
- , Yolanda Cañadas
- , Cristina Carbonell
- , Servando Cardona-Huerta
- , Antonio Augusto F. Carioca
- , Maria Sanchez Carpintero
- , Carlos Carpio Segura
- , Thássia M. T. Carratto
- , José Antonio Carrillo-Avila
- , Maria C. C. Carvalho
- , Carlos Casasnovas
- , Luis Castano
- , Carlos F. Castaño
- , Jose E. Castelao
- , Aranzazu Castellano Candalija
- , María A. Castillo
- , Francisco C. Ceballos
- , Jessica G. Chaux
- , Walter G. Chaves-Santiago
- , Sylena Chiquillo-Gómez
- , Marco A. Cid-Lopez
- , Oscar Cienfuegos-Jimenez
- , Rosa Conde-Vicente
- , M. Lourdes Cordero-Lorenzana
- , Dolores Corella

- , Almudena Corrales
- , Jose L. Cortes-Sanchez
- , Marta Corton
- , Tatiana X. Costa
- , Raquel Cruz
- , Marina S. Cruz
- , Luisa Cuesta
- , Gabriela C. R. Cunha
- , Gabriela V. da Silva
- , David Dalmau
- , Raquel C. S. Dantas-Komatsu
- , M. Teresa Darnaude
- , Raimundo de Andrés
- , Jéssica N. G. de Araújo
- , Carmen de Juan
- , Juan De la Cruz Troca
- , Carmen de la Horra
- , Ana B. de la Hoz
- , Alba De Martino-Rodríguez
- , Julianna Lys de Sousa Alves Neri
- , Victor del Campo-Pérez
- , Juan Delgado-Cuesta
- , Covadonga M. Diaz-Caneja
- , Anderson Díaz-Pérez
- , Aranzazu Diaz de Bustamante
- , Beatriz Dietl
- , Silvia Diz-de Almeida
- , Manoella do Monte Alves
- , Elena Domínguez-Garrido
- , Katiusse A. dos Santos
- , Alice M. Duarte
- , Jose Echave-Sustaeta
- , Rocío Eiros
- , César O. Enciso-Olivera
- , Gabriela Escudero
- , Pedro Pablo España
- , Gladys Mercedes Estigarribia Sanabria

- , María Carmen Fariñas
- , Marianne R. Fernandes
- , Ramón Fernández
- , Lidia Fernandez-Caballero
- , Ana Fernández-Cruz
- , María J. Fernandez-Nestosa
- , Uxía Fernández-Robelo
- , Amanda Fernández-Rodríguez
- , Marta Fernández-Sampedro
- , Ruth Fernández-Sánchez
- , Tania Fernández-Villa
- , Silvia Fernández Ferrero
- , Yolanda Fernández Martínez
- , Carmen Fernéndez Capitán
- , Patricia Flores-Pérez
- , Vicente Friaza
- , Lácides Fuenmayor-Hernández
- , Marta Fuertes Núñez
- , Victoria Fumadó
- , Ignacio Gadea
- , Lidia Gagliardi
- , Manuela Gago-Domínguez
- , Natalia Gallego
- , Cristina Galoppo
- , Inés García
- , Mercedes García
- , Leticia García
- , Carlos Garcia-Cerrada
- , Aitor García-de-Vicuña
- , Josefina Garcia-García
- , Irene García-García
- , Carmen García-Ibarbia
- , Andrés C. García-Montero
- , Ana García-Soidán
- , Elisa García-Vázquez
- , María Carmen García Torrejón
- , Emiliano Garza-Frias

- , Angela Gentile
- , Belén Gil-Fournier
- , Javier Gómez-Arrue
- , Mario Gómez-Duque
- , Luis Gómez Carrera
- , María Gómez García
- , Ángela Gómez Sacristán
- , Anna González-Neira
- , Javier González-Peñas
- , Manuel Gonzalez-Sagrado
- , Beatriz González Álvarez
- , Fernan Gonzalez Bernaldo de Quirós
- , Hugo Gonzalo Benito
- , Oscar Gorgojo-Galindo
- , Miguel Górgolas
- , Florencia Guaragna
- , Genilson P. Guegel
- , Beatriz Guillen-Guio
- , Encarna Guillen-Navarro
- , Pablo Guisado-Vasco
- , Juan F. Gutiérrez-Bautista
- , Luz D. Gutierrez-Castañeda
- , Sarah Heili-Frades
- , Estefania Hernandez
- , Luis D. Hernandez-Ortega
- , Guillermo Hernández-Pérez
- , Rebeca Hernández-Vaquero
- , Cristina Hernández Moro
- , Belen Herraez
- , M. Teresa Herranz
- , María Herrera
- , María José Herrero
- , Antonio Herrero-Gonzalez
- , Juan P. Horcajada
- , Natale Imaz-Ayo
- , Maider Intxausti-Urrutibeaskoa
- , María Íñiguez

- , Rafael H. Jacomo
- , Rubén Jara
- , Perez Maria Jazmin
- , Ángel Jiménez
- , Pilar Jiménez
- , Ignacio Jiménez-Alfaro
- , María A. Jimenez-Sousa
- , Iolanda Jordan
- , Rocío Laguna-Goya
- , Daniel Laorden
- , María Lasa-Lazaro
- , María Claudia Lattig
- , Ailen Lauriente
- , Anabel Liger Borja
- , Lucía Llanos
- , Amparo López-Bernús
- , Esther Lopez-Garcia
- , Rosario Lopez-Rodriguez
- , Miguel A. López-Ruz
- , Eduardo López Granados
- , Leonardo Lorente
- , José E. Lozano
- , María Lozano-Espinosa
- , Andre D. Luchessi
- , Ignacio Mahillo
- , Esther Mancebo
- , Carmen Mar
- , Cristina Marcelo Calvo
- , Miguel Marcos
- , Alba Marcos-Delgado
- , Alicia Marín Candon
- , Pablo Mariscal Aguilar
- , María M. Martín
- , María Dolores Martín
- , Vicente Martín
- , Marta Martin-Fernandez
- , Caridad Martín-López

- , José-Ángel Martín-Oterino
- , Laura Martin-Pedraza
- , María Martín-Vicente
- , Amalia Martinez
- , Ricardo Martínez
- , Juan José Martínez
- , Silvia Martínez
- , Eleno Martínez-Aquino
- , Óscar Martínez-González
- , Iciar Martinez-Lopez
- , Oscar Martinez-Nieto
- , Pedro Martinez-Paz
- , Angel Martinez-Perez
- , Andrea Martínez-Ramas
- , Michel F. Martinez-Resendez
- , Violeta Martínez Robles
- , Laura Marzal
- , Juliana F. Mazzeu
- , Jeane F. P. Medeiros
- , Kelliane A. Medeiros
- , Francisco J. Medrano
- , Xose M. Meijome
- , Natalia Mejuto-Montero
- , Ana Méndez-Echevarria
- , Humberto Mendoza Charris
- , Eleuterio Merayo Macías
- , Fátima Mercadillo
- , Arieh R. Mercado-Sesma
- , Pablo Minguez
- , Antonio J. J. Molina
- , Elena Molina-Roldán
- , Juan José Montoya
- , Vitor M. S. Moraes
- , Patricia Moreira-Escriche
- , Xenia Morelos-Arnedo
- , Antonio Moreno-Docón
- , Junior Moreno-Escalante

- , Victor Moreno Cuerda
- , Alberto Moreno Fernández
- , Rubén Morilla
- , Patricia Muñoz García
- , Pablo Neira
- , Julian Nevado
- , Israel Nieto-Gañán
- , Joana F. R. Nunes
- , Rocio Nuñez-Torres
- , Antònia Obrador-Hevia
- , J. Gonzalo Ocejo-Vinyals
- , Virginia Olivar
- , Silviene F. Oliveira
- , Lorena Ondo
- , Alberto Orfao
- , Luis Ortega
- , Eva Ortega-Paino
- , Fernando Ortiz-Flores
- , Rocio Ortiz-Lopez
- , José A. Oteo
- , Harry Pachajoa
- , Manuel Pacheco
- , Fredy Javier Pacheco-Miranda
- , Irene Padilla Conejo
- , Sonia Panadero-Fajardo
- , Mara Parellada
- , Roberto Pariente-Rodríguez
- , Estela Paz-Artal
- , Germán Peces-Barba
- , Miguel S. Pedromingo Kus
- , Celia Perales
- , Patricia Perez
- , César Pérez
- , Gustavo Perez-de-Nanclares
- , Felipe Pérez-García
- , Patricia Pérez-Matute
- , Alexandra Pérez-Serra

- , M. Elena Pérez-Tomás
- , Teresa Perucho
- , Lisbeth A. Pichardo
- , Susana M. T. Pinho
- , Mel·lina Pinsach-Abuin
- , Luz Adriana Pinzón
- , Guillermo Pita
- , Francesc Pla-Junca
- , Laura Planas-Serra
- , Ericka N. Pompa-Mera
- , Gloria L. Porras-Hurtado
- , Aurora Pujol
- , María Eugenia Quevedo Chávez
- , Maria Angeles Quijada
- , Inés Quintela
- , Diana Ramirez-Montaño
- , Soraya Ramiro León
- , Pedro Rascado Sedes
- , Delia Recalde
- , Emma Recio-Fernández
- , Salvador Resino
- , Adriana P. Ribeiro
- , Carlos S. Rivadeneira-Chamorro
- , Diana Roa-Agudelo
- , Montserrat Robelo Pardo
- , Marilyn Johanna Rodriguez
- , Fernando Rodriguez-Artalejo
- , Marena Rodríguez-Ferrer
- , Carlos Rodriguez-Gallego
- , José A. Rodriguez-Garcia
- , María A. Rodriguez-Hernandez
- , Antonio Rodriguez-Nicolas
- , Agustí Rodriguez-Palmero
- , Emilio Rodríguez-Ruiz
- , Paula A. Rodriguez-Urrego
- , Belén Rodríguez Maya
- , German Ezequiel Rodriguez Novoa

- , Federico Rojo
- , Andrea Romero-Coronado
- , Filomeno Rondón García
- , Lidia S. Rosa
- , Antonio Rosales-Castillo
- , Cladelis Rubio
- , María Rubio Olivera
- , Montserrat Ruiz
- , Francisco Ruiz-Cabello
- , Eva Ruiz-Casares
- , Juan J. Ruiz-Cubillan
- , Javier Ruiz-Hornillos
- , Pablo Ryan
- , Hector D. Salamanca
- , Lorena Salazar-García
- , Giorgina Gabriela Salgueiro Origlia
- , Pedro-Luis Sánchez
- , Clara Sánchez-Pablo
- , Olga Sánchez-Pernaute
- , Antonio J. Sánchez López
- , María Concepción Sánchez Prados
- , Javier Sánchez Real
- , Jorge Sánchez Redondo
- , Cristina Sancho-Sainz
- , Anna Sangil
- , Arnoldo Santos
- , Ney P. C. Santos
- , Agatha Schlüter
- , Sonia Segovia
- , Alex Serra-Llovich
- , Fernando Sevil Puras
- , Marta Sevilla Porras
- , Miguel A. Sicolo
- , Vivian N. Silbiger
- , Nayara S. Silva
- , Fabiola T. C. Silva
- , Cristina Silván Fuentes

- , Jordi Solé-Violán
- , José Manuel Soria
- , Jose V. Sorlí
- , Renata R. Sousa
- , Juan Carlos Souto
- , Karla S. C. Souza
- , Vanessa S. Souza
- , John J. Sprockel
- , José Javier Suárez-Rama
- , David A. Suarez-Zamora
- , Xiana Taboada-Fraga
- , Eduardo Tamayo
- , Alvaro Tamayo-Velasco
- , Juan Carlos Taracido-Fernandez
- , Nathali A. C. Tavares
- , Carlos Tellería
- , Jair Antonio Tenorio Castaño
- , Alejandro Teper
- , Juan Torres-Macho
- , Lilian Torres-Tobar
- , Ronald P. Torres Gutiérrez
- , Jesús Troya
- , Miguel Urioste
- , Juan Valencia-Ramos
- , Agustín Valido
- , Juan Pablo Vargas Gallo
- , Belén Varón
- , Romero H. T. Vasconcelos
- , Tomas Vega
- , Santiago Velasco-Quirce
- , Valentina Vélez-Santamaría
- , Virginia Víctor
- , Julia Vidán Estévez
- , Miriam Vieitez-Santiago
- , Carlos Vilches
- , Lavinia Villalobos
- , Felipe Villar

- , Judit Villar-Garcia
- , Cristina Villaverde
- , Pablo Villoslada-Blanco
- , Ana Virseda-Berdices
- , Zuleima Yáñez
- , Antonio Zapatero-Gaviria
- , Ruth Zarate
- , Sandra Zazo
- , Miguel López de Heredia
- , Ingrid Mendes
- , Rocío Moreno
- , Esther Sande
- , Carlos Flores
- , José A. Riancho
- , Augusto Rojas-Martinez
- , Pablo Lapunzina
- & Angel Carracedo

ISARICC Investigators

- **Co-Investigators**
  - J. Kenneth Baillie
  - , Peter J. M. Openshaw
  - , Malcolm G. Semple
  - , Beatrice Alex
  - , Petros Andrikopoulos
  - , Benjamin Bach
  - , Wendy S. Barclay
  - , Debby Bogaert
  - , Meera Chand
  - , Kanta Chechi
  - , Graham S. Cooke
  - , Ana da Silva Filipe
  - , Thushan de Silva
  - , Annemarie B. Docherty
  - , Gonçalo dos Santos Correia

- , Marc-Emmanuel Dumas
- , Jake Dunning
- , Tom Fletcher
- , Christopher A. Green
- , William Greenhalf
- , Julian Griffin
- , Rishi K. Gupta
- , Ewen M. Harrison
- , Antonia Y. W. Ho
- , Karl Holden
- , Peter W. Horby
- , Samreen Ijaz
- , Say Khoo
- , Paul Klenerman
- , Andrew Law
- , Matthew Lewis
- , Sonia Liggi
- , Wei Shen Lim
- , Lynn Maslen
- , Alexander J. Mentzer
- , Laura Merson
- , Alison M. Meynert
- , Shona C. Moore
- , Mahdad Noursadeghi
- , Michael Olanipekun
- , Anthonia Osagie
- , Massimo Palmarini
- , Carlo Palmieri
- , William A. Paxton
- , Georgios Pollakis
- , Nicholas Price
- , Andrew Rambaut
- , David L. Robertson
- , Clark D. Russell
- , Vanessa Sancho-Shimizu
- , Caroline Sands
- , Janet T. Scott

    - , Louise Sigfrid
    - , Tom Solomon
    - , Shiranee Sriskandan
    - , David Stuart
    - , Charlotte Summers
    - , Olivia V. Swann
    - , Zoltan Takats
    - , Panteleimon Takis
    - , Richard S. Tedder
    - , A. A. Roger Thompson
    - , Emma C. Thomson
    - , Ryan S. Thwaites
    - , Lance C. W. Turtle
    - , Maria Zambon
    - & Gail Carson

- **Data Analysis Team**

    - Thomas M. Drake
    - , Cameron J. Fairfield
    - , Stephen R. Knight
    - , Kenneth A. Mclean
    - , Derek Murphy
    - , Lisa Norman
    - , Riinu Pius
    - & Catherine A. Shaw

- **Data Architecture Team**

    - Marie Connor
    - , Jo Dalton
    - , Carrol Gamble
    - , Michelle Girvan
    - , Sophie Halpin
    - , Janet Harrison
    - , Clare Jackson

    - , Laura Marsh
    - , Stephanie Roberts
    - , Egle Saviciute
    - , Sara Clohisey
    - , Ross Hendry
    - , Susan Knight
    - , Eva Lahnsteiner
    - , Andrew Law
    - , Gary Leeming
    - , Lucy Norris
    - , James Scott-Brown
    - , Sarah Tait
    - & Murray Wham

- **Data Analysis and Management Team**

    - James Lee
    - , Daniel Plotkin
    - & Seán Keating

- **Project Administration Team**

    - Cara Donegan
    - & Rebecca G. Spencer

- **Project Management Team**

    - Chloe Donohue
    - , Fiona Griffiths
    - , Hayley Hardwick
    - & Wilna Oosthuyzen

The 23andMe COVID-19 Team

- Janie F. Shelton
- , Anjali J. Shastri

- , Chelsea Ye
- , Catherine H. Weldon
- , Teresa Filshtein-Sonmez
- , Daniella Coker
- , Antony Symons
- , Jorge Esparza-Gordillo
- , Stella Aslibekyan
- & Adam Auton

Contributions

E.P.-C., K. Rawlik, K.M., S.K., C.P.P., J.F.W., V.V., M.A., A.D.L., E.J.P., R.C., A.C., A.F., L.M., K. Rowan, A.C.P., A.L., S.C.H. and J.K.B. contributed to design. E.P.-C., K. Rawlik, A.D.B., T.Q., Y.W., I.N., G.A.M., M.Z., L.K., A.K., A.R., T.M., J.Y., A.L., B.F., S.C.H. and J.K.B. contributed to data analysis. E.P.-C., K. Rawlik, I.N., A.K., A.R., J.M., C.D.R., A.L., B.F. and S.C.H. contributed to bioinformatics. E.P.-C., K. Rawlik, I.N., G.A.M., M.Z., A.K., J.M., C.D.R., R.T., D. McAuley, A.N., M.G.S., B.F., S.C.H. and J.K.B. contributed to writing and reviewing the manuscript. I.N., F.G., W.O., K.M., S.K., D. Maslove, A.N., M.G.S., J.K., M.S.-H., C.S., C.H., P.H., L.L., D. McAuley, H.M., P.J.M.O., C.B., T.W., A.T., C.F., J.A.R., A.R.-M., P.L., C.P.P., A.F., L.M., K. Rowan, A.L., B.F. and S.C.H. contributed to oversight. F.G. and W.O. contributed to project management. F.G., W.O. and J.K.B. contributed to ethics and governance. K.M., A.F. and L.M. contributed to sample handling and sequencing. C.P.P., K. Rowan, S.C.H. and J.K.B. contributed to conception. C.P.P., J.F.W., V.V., M.A., A.D.L., E.J.P., R.C., A.C., K. Rowan and A.C.P. contributed to reviewing the manuscript. K. Rowan and A.L. contributed to clinical data management. J.K.B. contributed to scientific leadership.

Corresponding author

Correspondence to J. Kenneth Baillie.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks Jacques Fellay and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Pipeline of meta-analysis and post-GWAS analyses.

Red border indicates that the data is only available for the hospitalized phenotype, while a black border indicates that the analysis was performed for the critical illness phenotype.

Extended Data Fig. 2 Miami plots.

Meta-analysis results are shown for a) critical and b) hospitalized phenotypes. In each plot results obtained using all cohorts are shown at the top and using GenOMICC data only at the bottom. Independent lead variants in the analyses of all cohorts are annotated with associated genes. Genome-wide significant associations that have not been previously reported are indicated in bold.

Extended Data Fig. 3 Comparison of effect size estimates.

GenOMICC is compared with the critical and hospitalized phenotype definitions in the SCOURGE, 23andMe, and HGI analyses. The black line indicates the best linear fit, given by the equation in each plot, obtained using Orthogonal Distance Regression to account for estimation errors in both sets of effects in the comparison.

Extended Data Fig. 4 Study weightings for (a) critical and (b) hospitalized COVID-19.

Mean +/− standard deviation of weights assigned to each data source in meta-analyses for all significant SNPs.

Extended Data Fig. 5 Cartoon showing postulated roles for genes and mediators implicated in the pathogenesis of critical COVID-19 by GenOMICC GWAS, TWAS and Mendelian randomization.

Postulated roles for genetic variants are shown in a highly simplified format to illustrate potential roles in pathogenesis, with the shaded background indicating the hypothetical impact of the host immune response over time[17]. Host immune processes are divided into those that are thought to play a role in controlling viral replication early in disease (orange section, showing "adaptive" response), and those implicated in driving hypoxaemic respiratory failure later in disease (green section, showing "maladaptive" response). Bold type gene names indicate a higher level of confidence in both the gene identification and the biological role.

Extended Data Fig. 6 Functional genomics analyses for TYK2.

(a) Effect size plot for effect of multiple variants on TYK2 expression (eQTLgen, x-axis) against increasing susceptibility to critical COVID-19 (β_{xy} = 0.53; $P_{xy}=1.2\times 10^{-23}$). Colour shows linkage disequilibrium (LD) with the missense variant rs34536443. (b) Crystal structure of TYK2 kinase domain (Protein Data Bank ID

4GVJ[39]) in two views that differ by a 45° rotation around a horizontal axis. The side chain of P1104 is shown as connected spheres with a nitrogen atom coloured in blue. Carbon, oxygen, nitrogen and phosphorus atoms of ATP are shown as magenta, red, blue and orange connected spheres, respectively. The N-terminal region of the kinase domain is not shown in the second view for clarity. The right-most panel shows a close view of P1104 and neighbouring residues with their side chains shown as sticks. Numbering of residues corresponds to UniProtKB entry P29597. P1104 is in the catalytic kinase domain and proximal to the ATP-binding site; TYK2 P1104A is catalytically impaired[40].

Extended Data Fig. 7 Steroid treatment and vaccination status.

Data are shown for a subset of GenOMICC cases who were also recruited to the ISARIC4C study in the UK.

Extended Data Table 1 pQTL GSMR results table
Extended Data Table 2 Replication table

Supplementary information

Supplementary Information

Supplementary Sections 1–13, including Supplementary Figs. 1–54 and Supplementary Tables 7–17.

Reporting Summary

Peer Review File

Supplementary Table 1

Description of the cohorts used in critical and hospitalized meta-analyses. Cohorts are divided by ancestry and genotyping method (whole-genome sequencing or microarray genotyping). In cohorts in which data are available, the median age with s.d. in parentheses, percentage of female

cases, number of female and male cases and controls is shown. NA, data are not available for the cohort. Country of origin indicates the country in which individuals in the cohort were recruited; GenotypingPlatform indicates the array or WGS platform used for genotyping; reference indicates the reference of the publication (if the data have already been published). 'In GenOMICC v2' is an indication of whether the dataset was included in a previous GenOMICC paper[1].

Supplementary Table 2

Full results for colocalization and TWAS analyses in lungs, blood, monocytes and across multiple tissue types (metaTWAS). Colocalization results are reported between significant TWAS genes and eQTLs in GTExv8 in lungs, blood and monocytes, and in eqtlGEN. rsid indicates chromosome, position, reference and alternative alleles; gene.tested is the significant TWAS gene, and ensembl.id is the Ensembl ID corresponding to the significant gene. PP.H3.1e-5 is the posterior probability of independent signals with a prior for colocalization of 5×10^{-5}, PP.H4.1e-5 is the posterior probability of colocalization with a prior of 5×10^{-5}, PP.H3.5e-5 is the posterior probability of independent signals with a prior for colocalization of 5×10^{-5} and PP.H4.5e-5 is the probability of colocalization with a prior of 5×10^{-5}. Colocalization was considered to be significant when PP.H3 was lower than 0.5 and PP.H4 was the highest posterior probability. Significant genes after Bonferroni correction in a TWAS meta-analysis of lungs, blood, monocytes and all tissues in GTExv8 and the 'all critical cohorts' GWAS. Ensembl ID, gene name, *P* value of the meta-analysis, number of SNPs used in to model gene expression in all tissues, mean *z* score for all tissues and its s.d. are shown.

Supplementary Table 3

The full results from GSMR analysis for protein level. Exposure indicates the protein name used as exposure for the analysis, bxy is the effect size, se is the standard error, p is the *P* value of the analysis, nsnp is the number of SNPs used as instruments and multi_snp_based_heidi_outlier is the *P* value of the Heidi test.

Supplementary Table 4

Full results from GSMR analysis for RNA-seq data from eQTLGEN. Exposure indicates the gene name used as exposure for the analysis, bxy is the effect size, se is the standard error, p is the *P* value of the analysis, nsnp is the number of SNPs used as instruments and multi_snp_based_heidi_outlier is the *P* value of the Heidi test.

Supplementary Table 5

Table of variants in credible sets with 95% probability of containing the causal SNP. Credible set ID is the credible set index to which the variant belongs; posterior is the posterior probability of causality for the variant. Variant indicates chromosome, position, reference and alternative alleles; beta is the effect, beta.se is the error and P is the *P* value.

Supplementary Table 6

The full results of the gene-level analysis performed using the mBAT-combo method. Gene indicates ensembl ID, gene_name indicates name of the gene, gene_p is the *P* value of the gene-level test, nsps is the number of SNPs used for the test, lead_snp is the chromosome, position, reference and alternative alleles for the SNPs with lowest *P* value in the region, and lead_snp_p indicates the *P* value of this lead SNP.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted

by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Pairo-Castineira, E., Rawlik, K., Bretherick, A.D. *et al.* GWAS and meta-analysis identifies 49 genetic variants underlying critical COVID-19. *Nature* **617**, 764–768 (2023). https://doi.org/10.1038/s41586-023-06034-3

- Received: 22 November 2022
- Accepted: 27 March 2023
- Published: 17 May 2023
- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-06034-3

Share this article

Anyone you share the following link with will be able to read this content:

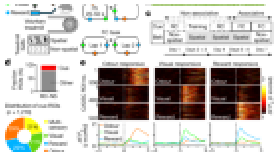
Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

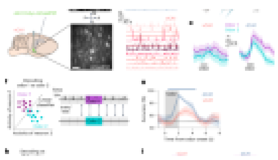
This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06034-3



Adaptive stimulus selection for consolidation in the hippocampus

08 December 2021

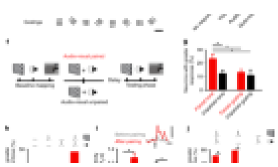
Satoshi Terada, Tristan Geiller, … Attila Losonczy



Neural dynamics underlying associative learning in the dorsal and ventral hippocampus

03 April 2023

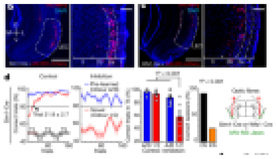
Jeremy S. Biane, Max A. Ladow, … Mazen A. Kheirbek



Audio-visual experience strengthens multisensory assemblies in adult mouse visual cortex

12 December 2019

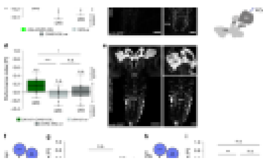
Thomas Knöpfel, Yann Sweeney, … Samuel J. Barnes



[Dopamine facilitates associative memory encoding in the entorhinal cortex]

22 September 2021

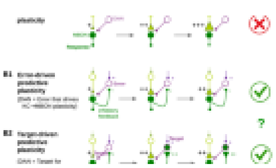
Jason Y. Lee, Heechul Jun, … Kei M. Igarashi



[Reward signaling in a recurrent circuit of dopaminergic neurons and peptidergic Kenyon cells]

15 July 2019

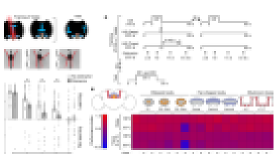
Radostina Lyutova, Mareike Selcho, … Dennis Pauls



[Predictive olfactory learning in Drosophila]

24 March 2021

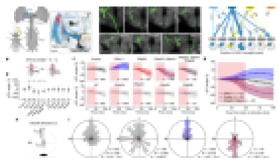
Chang Zhao, Yves F. Widmer, … Walter Senn



[Differential mechanisms underlie trace and delay conditioning in Drosophila]

16 February 2022

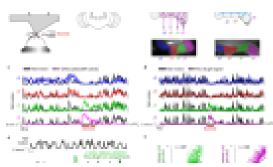
Dhruv Grover, Jen-Yung Chen, … Ralph J. Greenspan



[Synaptic gradients transform object location to action]

04 January 2023

Mark Dombrovski, Martin Y. Peek, … Gwyneth M. Card



[Context-dependent representations of movement in Drosophila dopaminergic reinforcement pathways]

25 October 2021

Aryeh Zolin, Raphael Cohn, … Vanessa Ruta

- Article
- Open Access
- Published: 26 April 2023

Multisensory learning binds neurons into a cross-modal memory

engram

- Zeynep Okray ,
- Pedro F. Jacob,
- Ciara Stern,
- Kieran Desmond,
- Nils Otto,
- Clifford B. Talbot,
- Paola Vargas-Gutierrez &
- …
- Scott Waddell

Show authors

Nature **volume 617**, pages 777–784 (2023)

- 12k Accesses

- 95 Altmetric

- Metrics details

Abstract

Associating multiple sensory cues with objects and experience is a fundamental brain process that improves object recognition and memory performance. However, neural mechanisms that bind sensory features during learning and augment memory expression are unknown. Here we demonstrate multisensory appetitive and aversive memory in *Drosophila*. Combining colours and odours improved memory performance, even when each sensory modality was tested alone. Temporal control of neuronal function revealed visually selective mushroom body Kenyon cells (KCs) to be required for enhancement of both visual and olfactory memory after multisensory training. Voltage imaging in head-fixed flies showed that multisensory learning binds activity between streams of modality-specific KCs so that unimodal sensory input generates a multimodal neuronal

response. Binding occurs between regions of the olfactory and visual KC axons, which receive valence-relevant dopaminergic reinforcement, and is propagated downstream. Dopamine locally releases GABAergic inhibition to permit specific microcircuits within KC-spanning serotonergic neurons to function as an excitatory bridge between the previously 'modality-selective' KC streams. Cross-modal binding thereby expands the KCs representing the memory engram for each modality into those representing the other. This broadening of the engram improves memory performance after multisensory learning and permits a single sensory feature to retrieve the memory of the multimodal experience.

Main

Life is a rich multisensory experience for most animals. As a result, nervous systems have evolved to use multisensory representations of objects, scenes and events to most effectively guide behaviour[1]. It is widely appreciated that multisensory learning improves memory performance, from children in the classroom to rodents and insects in controlled laboratory experiments[2,3,4,5]. Moreover, an apparently universal and unexplained feature of multisensory learning is that it improves subsequent memory performance even for the separate unisensory components[3,5]. Studies in humans and other mammals have suggested that multisensory learning benefits from interactions between modality-specific cortices that were co-active during training, and that individual senses can reactivate both areas at testing[3,6,7,8,9]. In addition, cells in different brain regions respond to multiple sensory cues and the proportions or numbers change after multisensory learning[1,10,11,12]. However, we currently lack detailed mechanistic understanding of how multisensory learning converts neurons from being modality selective to multimodal, and how enhanced multisensory and unisensory memory performance could be supported by such a process.

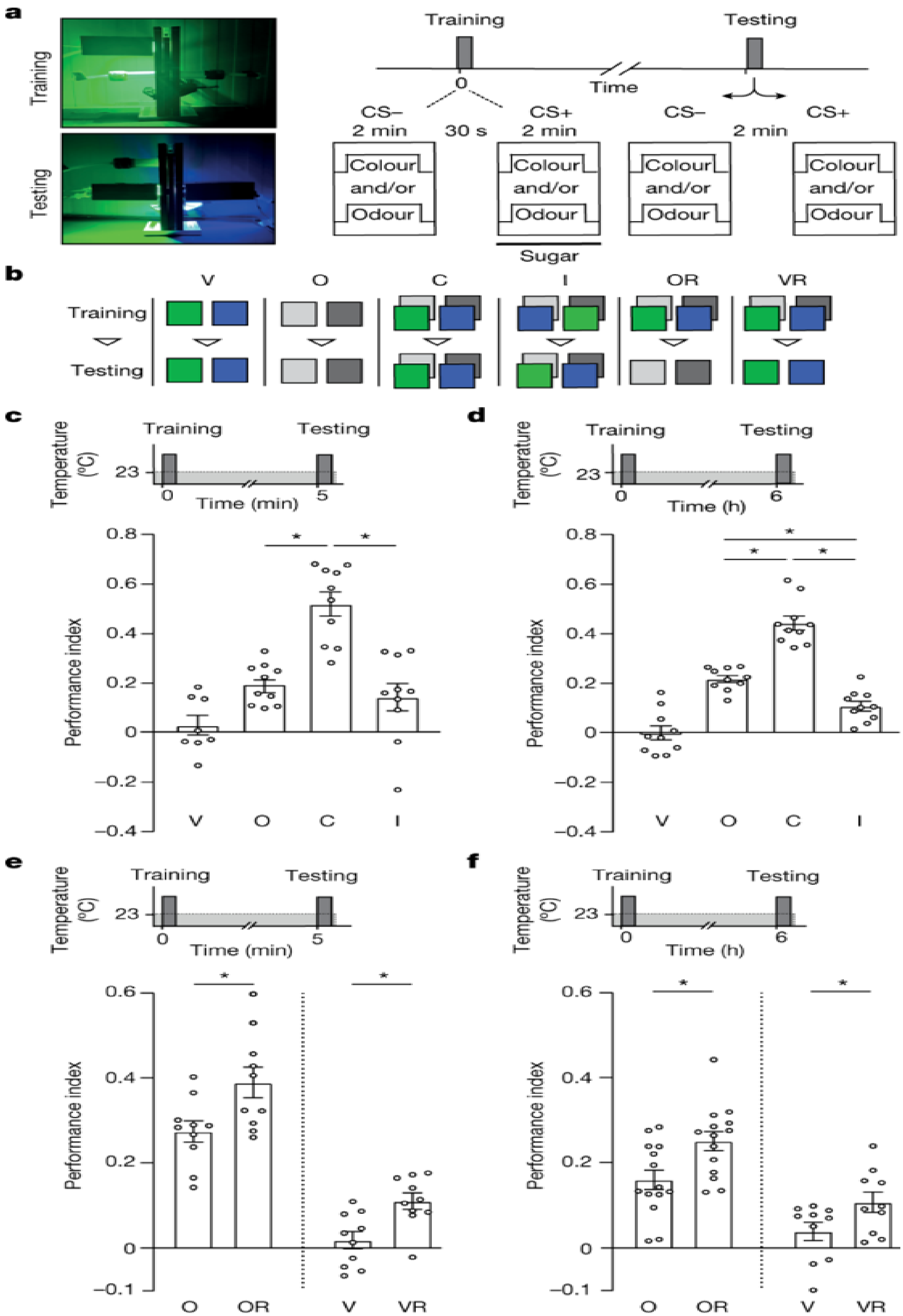
In *Drosophila*, unique populations of mushroom body (MB) KCs receive predominant and anatomically segregated dendritic input from olfactory or visual projection neurons (as well as local visual interneurons) and their axons project as parallel streams into the MB lobes. There, successive compartments of the axonal arbor of each KC are intersected by the presynapses of dopaminergic neurons (DANs) that convey the reinforcing

effects of appetitive or aversive stimuli[13,14]. Reinforcing dopamine depresses synapses between active KCs and the compartment-restricted dendrites of downstream MB output neurons to code valence-specific memories[15,16,17].

Multisensory learning improves memory

To study multisensory learning in *Drosophila*, we adapted the olfactory T-maze[18] so that colours and odours can be presented together (Fig. 1a). Food-deprived flies were trained by presenting them with a colour and/or odour (conditioned stimulus minus (CS−)), followed by another colour and/or odour (conditioned stimulus plus (CS+)) paired with a sugar reward (Fig. 1a,b). When trained and tested with only colours (visual learning), flies did not show significant learned preference for the previously sugar-paired colour (Fig. 1b–d and Extended Data Fig. 1a). However, combining colours with odours (congruent protocol) produced robust and long-lasting memory, which was significantly enhanced over that formed by training with only odours (olfactory learning) (Fig. 1b–d and Extended Data Fig. 1a,b). If colour and odour combinations were swapped between training and testing (incongruent protocol) (Fig. 1b–d and Extended Data Fig. 1a,b), no memory enhancement was observed. Furthermore, memory enhancement was not apparent if the same colour was presented with CS− and CS+ odours during training and testing (Extended Data Fig. 1c). The memory-enhancing effect of multisensory learning therefore requires a learned relationship between specific colour and odour combinations. For memory measured 6 h after training, the incongruent protocol revealed significantly decreased performance compared with that following olfactory learning (Fig. 1d), suggesting that flies are conflicted when colour–odour contingency is switched between training and testing.

Fig. 1: Multisensory learning enhances memory performance.



a
Training
Testing
Training
Testing
0
Time
CS–
2 min
30 s
CS+
2 min
CS–
2 min
CS+
Colour and/or Odour
Sugar
b
V
O
C
I
OR
VR
Training
Testing
c
Temperature (°C)
23
Training
Testing
0
Time (min)
5
Performance index
V
O
C
I
d
Temperature (°C)
Training
Testing
Time (h)
6
Performance index
e
Training
Testing
Time (min)
Performance index
O
OR
V
VR
f
Training
Testing
Time (h)
Performance index
O
OR
V
VR

a, Apparatus for multisensory training and testing (left), and the experimental timeline (right). **b**, Protocols. The green and blue squares represent colours, and the light and dark grey squares represent 3-octanol (OCT) and 4-methylcyclohexanol (MCH) odours. For visual (V) learning, colours were used as CS+ and CS−. For olfactory (O) learning, odours were used as CS+ and CS−. For the congruent (C) protocol, colours + odours were combined as CS+ and CS− and the same colour + odour combinations were used for training and testing. For the incongruent (I) protocol, colours + odours were combined as CS+ and CS−, but combinations were switched between training and testing. For olfactory retrieval (OR), colour + odour combinations were used for training but only odours were used for testing. For visual retrieval (VR), colour + odour combinations were used for training but only colours were used for testing. **c**,**d**, Training and testing timelines (top), and immediate (**c**) and 6 h (**d**) memory for V, O, C and I protocols (bottom). **e**,**f**, Timelines (top), and multisensory training with colours + odours tested immediately (**e**) and 6 h (**f**) after training for each individual modality (bottom). Asterisks denote significant difference ($P < 0.05$). Data are presented as mean ± s.e.m. Individual data points displayed as dots correspond to independent experiments. Groups were compared using one-way analysis of variance (ANOVA) with Tukey's test (**c**,**d**) and unpaired two-sided *t*-test (**e**,**f**); exact *P* values and comparisons are provided in Supplementary Information. $n = 8$ for V and $n = 10$ for O, C and I (**c**); $n = 10$ (**d**); $n = 10$ (**e**); and $n = 14$ for O and OR and $n = 10$ for V and VR (**f**).

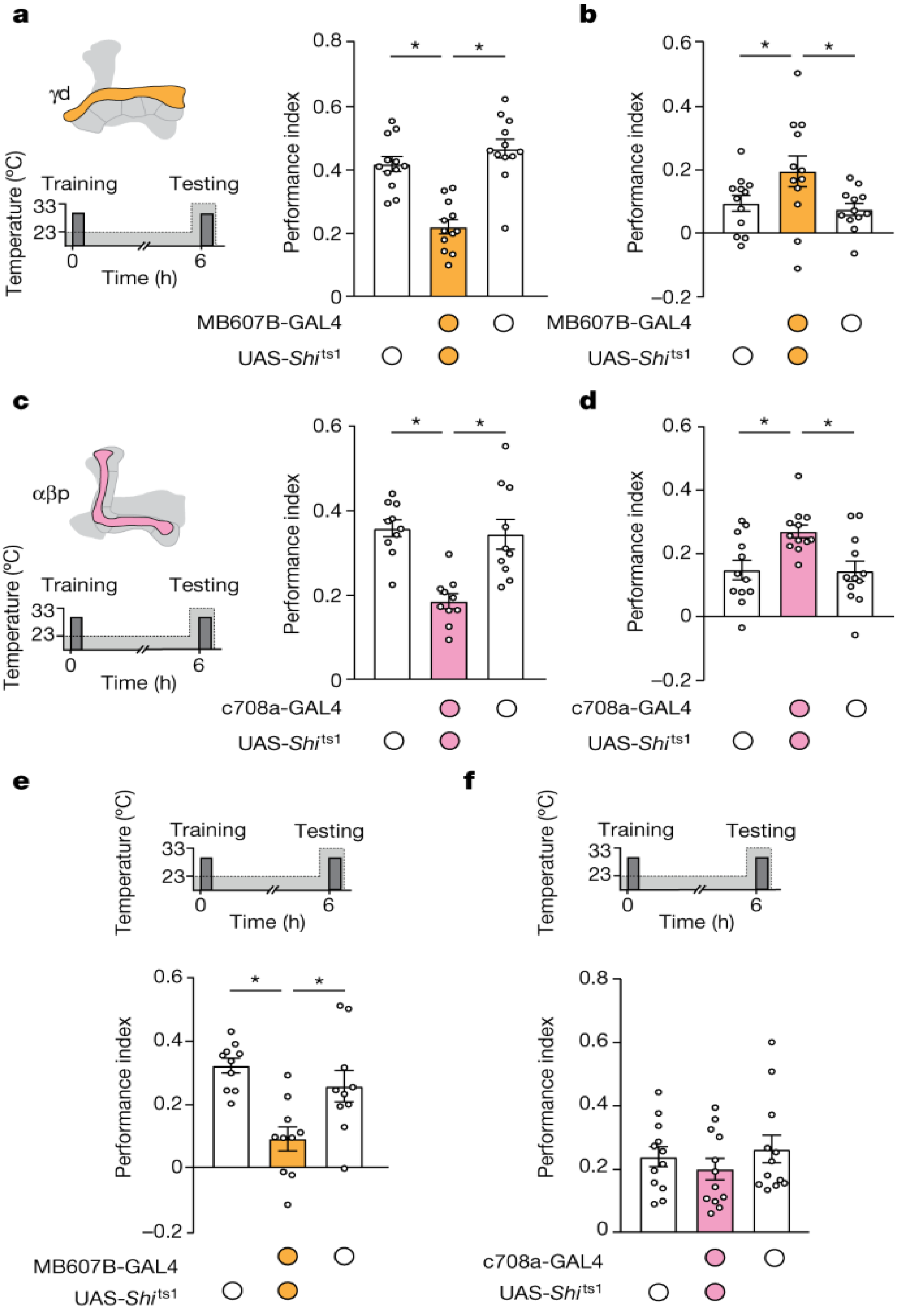
To further investigate multimodal memory enhancement, we restricted presentation of multisensory cues to either training or testing. Multisensory training improved memory retrieval even when each modality was presented alone during testing (olfactory retrieval and visual retrieval) (Fig. 1e,f). By contrast, presenting multisensory stimuli only during testing did not facilitate performance (multisensory retrieval) (Extended Data Fig. 1d). Moreover, the greatest improvement in performance was observed when multisensory stimuli were used during training and testing (Extended Data Fig. 1e). Therefore, multisensory training enhances memory performance for the individual colour and odour memory components, and congruence of colour and odour information between training and testing further improves performance. Although our experiments and elsewhere[19] imply that flies

distinguish green and blue colours, we do not discount a contribution of hue and luminance.

Olfactory retrieval requires visual KCs

Dendrites of the numerically larger populations of olfactory KCs within the αβ-lobe, α′β′-lobe and γ-lobe (that is, γ-main (γm KCs)) occupy the main calyces of the MB, whereas the relatively small populations of αβ-posterior (αβp) and γ-dorsal (γd) KCs receive predominantly visual information via dendrites in the accessory calyces[14]. γd KCs were previously implicated in colour learning[19,20]. We tested roles for visual γd and αβp KCs in multisensory learning and memory, using cell-specific expression of a UAS-*Shibire*^{ts1} (*Shi*^{ts1}) transgene, which encodes a dominant temperature-sensitive dynamin[21]. At temperatures over 30 °C, *Shi*^{ts1} blocks membrane recycling and thus impairs synaptic transmission, whereas function can be restored by returning flies to less than 29 °C. Blocking output from γd and αβp KCs during testing at 6 h abolished visual enhancement of performance in the congruent protocol and removed the interference of incongruence; in both instances, memory performance was similar to that of flies tested with odours alone (Fig. 2a–d and Extended Data Fig. 2a–f). These results suggest that activity in γd and αβp KCs represents the visual component of multisensory memory (see also Extended Data Fig. 2g,h). Blocking γd KCs (but not αβp KCs) output during testing also impaired performance for odour-only memory retrieval in multisensory trained flies (Fig. 2e,f), despite having no effect on memory retrieval in flies trained with only odours[22] (Extended Data Fig. 2a,d). This unexpected result led us to hypothesize that multisensory learning might expand the representation of odours to include 'visual' γd KCs.

Fig. 2: Enhanced performance following multisensory learning requires visually responsive γd and $\alpha\beta_p$ KCs.


a
γd
Training
Testing
Temperature (°C)
33
23
0
6
Time (h)
Performance index
MB607B-GAL4
UAS-Shi ts1
b
c
αβp
c708a-GAL4
d
e
f

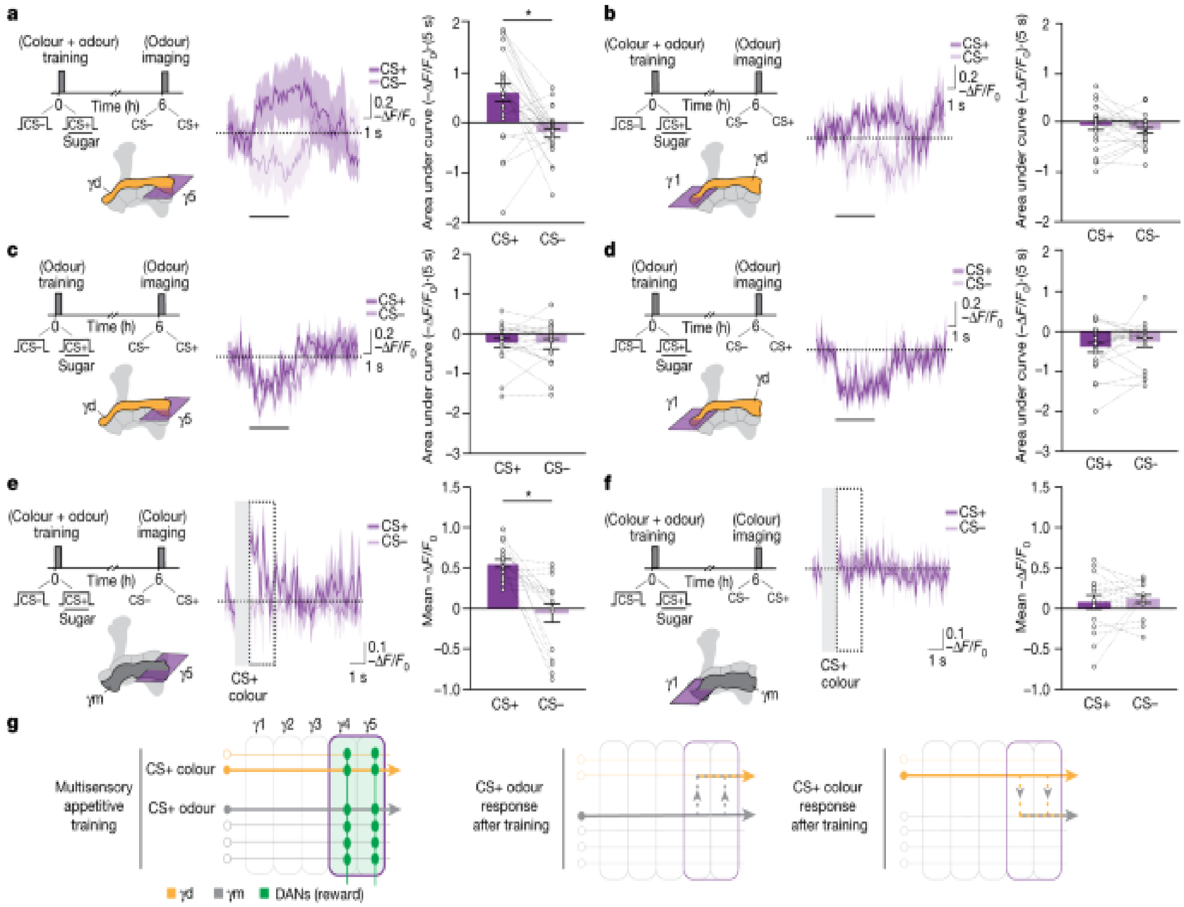
a, Schematic of γd KCs (top left), and the timeline with temperature shifting (dashed line) (bottom left). Blocking output of γd KCs during testing using MB607B-GAL4;UAS-Shi^{ts1} in the congruent protocol is also shown (right). **b**, Blocking output of γd KCs during testing in the incongruent protocol. **c**, Schematic of $\alpha\beta_p$ KCs (top left), and the timeline with temperature shifting (bottom left). Blocking output of $\alpha\beta_p$ KCs during testing using c708a-GAL4;UAS-Shi^{ts1} in the congruent protocol is also shown (right). **d**, Blocking output of $\alpha\beta_p$ KCs during the incongruent protocol. **e**,**f**, Timeline with temperature shifting (top), and blocking output of γd (**e**) and $\alpha\beta_p$ (**f**) KCs during testing of olfactory retrieval of multisensory memory (bottom). Asterisks denote significant difference ($P < 0.05$). Data are presented as mean ± s.e.m. Individual data points displayed as dots correspond to independent experiments. All groups were compared using one-way ANOVA with Tukey's test; exact P values and comparisons are provided in Supplementary Information. $n = 12$ (**a**,**b**,**d**,**f**) and $n = 10$ (**c**,**e**). See Extended Data Fig. 2 for controls.

Neurons gain cross-modal activation

To directly test for learned odour-evoked responses in γd KCs after multisensory training, we expressed the voltage sensor UAS-ASAP2f[23] in γd KCs and performed two-photon functional imaging (Fig. 3a–d and Extended Data Fig. 3a,b,f,g). Flies received multisensory (colour + odour), unisensory (odour) or unpaired (sugar presented 2 min after colour + odour) training. Six hours after training, flies were imaged for CS+ and CS− odour responses. Recordings of γd KC axons were made in the terminal γ5 compartment of the MB horizontal lobe, as sugar-rewarding DANs drive learning-relevant presynaptic depression of KC–MB output neuron synapses in γ4 and γ5 compartments[24,25]. For comparison, we imaged the responses of γd KCs in the proximal γ1 compartment, which houses the presynaptic field of DANs providing aversive teaching signals[26,27,28]. Odour presentation was previously shown to evoke slow inhibition in γd[19] and $\alpha\beta_p$[22] KCs of naive flies. We found that presentation of the CS− odour evoked hyperpolarization of γd KCs in both γ1 and γ5 compartments, regardless of the training protocol (Fig. 3a–d and Extended Data Fig. 3a,b, light purple trace). However, after multisensory training, the CS+ odour produced significant

depolarization of γd KC axons in the γ5 compartment (Fig. 3a, dark purple trace). CS+ responses in the γ1 compartment appeared less inhibitory than those to CS− (Fig. 3b; although the responses were statistically indistinguishable), perhaps due to the PPL1-γ1pedc DANs being modulated by the hunger state of the fly[29,30]. The multisensory training-driven sign reversal of the γd KC odour response in γ5 did not occur following unisensory odour-only training (Fig. 3c, dark purple trace) or unpaired training (Extended Data Fig. 3a, dark purple trace). In these cases, both CS+ and CS− odours evoked hyperpolarization in γ1 and γ5 compartments (Fig. 3c,d and Extended Data Fig. 3a,b). Imaging colour-evoked signals revealed strong responses to both colours in γd KCs of naive flies (Extended Data Fig. 3c). Recording γm KCs revealed a pronounced γ5 compartment-restricted gain of excitation by the CS+ colour after multisensory training, with no alteration of responses in the γ1 compartment (Fig. 3e,f; note that laser-scanning image detection is blocked by a shutter during colour presentation). Pulsing coloured light that is required for these imaging experiments did not impair multisensory learning and visual retrieval (Extended Data Fig. 3d,e) or γm KC responses to odours (Extended Data Fig. 3h,i). These results indicate that dopaminergic reward teaching signals broaden CS+ odour-evoked and CS+ colour-evoked excitation within the γ-KC ensemble by recruiting the γ5 segments of γd KC axons to be odour activated and γm colour activated (Fig. 3g). These larger colour and odour memory engrams provide a mechanism for how odour and colour memory performance is enhanced following multisensory training (Fig.1e,f), and explains why odour memory retrieval in this context acquires a requirement for γd KC output (Fig. 2e).

Fig. 3: Multisensory learning converts γd KCs to be odour activated and γm KCs colour activated.



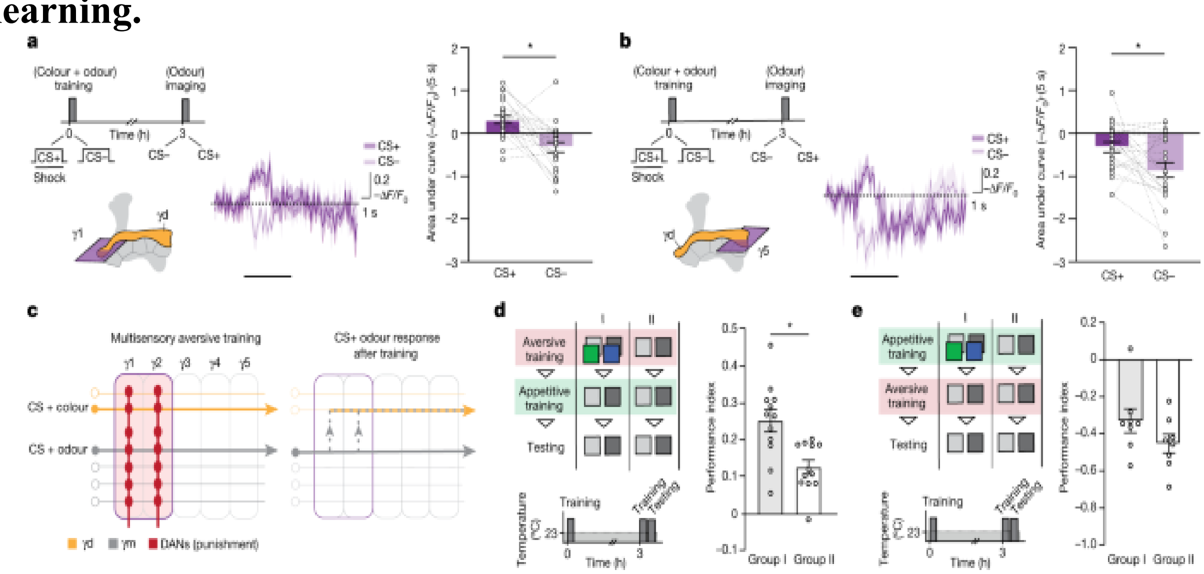
In panels **a**–**f**, timelines for appetitive multisensory (colour + odour) or appetitive unisensory (odour) training followed by odour or colour response imaging (top left); imaging plane in either the γ5 or γ1 region of γd or γm KC axons (lower left); traces of CS+ and CS− odour-evoked activity (middle); and quantification of responses (right) are shown. **a**, γ5 Region of γd KC axons showed an excitatory response to the CS+ odour (a decrease in fluorescence increases the $-\Delta F/F_0$ of the ASAP2f voltage sensor) after multisensory training. γ5 Axons were inhibited by the CS− odour (decrease in $-\Delta F/F_0$). **b**, Excitatory responses to the CS+ odour were not observed in γ1 and the CS− odour elicited inhibition. **c**,**d**, Both γ5 (**c**) and γ1 (**d**) regions showed inhibition to CS+ and CS− odours after appetitive unisensory (odour) training. **e**, γ5 Region of γm KC axons showed an excitatory response only to the CS+ colour after multisensory training. **f**, Excitation to the CS+ colour was not observed in γ1. For all traces and quantifications in this study, CS+ data correspond to average responses in which half of the trials used blue or MCH as CS+ and the other half used green or OCT as CS+. The same applies for CS− data. Odour-evoked or colour-evoked

activity traces show mean (solid line) with s.e.m. (shadow). Horizontal dashed lines indicates the baseline activity. In **a**–**d**, the solid black line below traces marks 5-s odour exposure. In **e** and **f**, the vertical grey bar corresponds to the 0.75-s colour presentation when image acquisition is shuttered and dotted box corresponds to the 1.75-s period of quantification. Asterisks denote significant difference between averaged CS+ and CS− responses ($P < 0.05$). CS+ and CS− responses for each fly are connected by a dashed line. All groups were compared using paired two-sided *t*-test; exact *P* values and comparisons are provided in Supplementary Information. n = 26 flies (**a**,**b**); n = 24 flies (**c**); n = 22 flies (**d**); and n = 16 flies (**e**,**f**). **g**, MB model for appetitive multisensory colour + odour training followed by unisensory odour or colour testing. γm KCs receive dendritic olfactory input and γd visual input. Both γ-KC types project axons through γ1–γ5 compartments of the MB γ-lobe. Appetitive training (left) engages reward DANs (green) innervating γ4 and γ5, whose released dopamine encodes learning by depressing synapses[49] between odour-activated KCs and avoidance-directing MB output neurons (not illustrated)[14]. Dopamine signalling during multisensory learning also binds γm and γd KC activity in γ4–γ5 compartments. During future unisensory odour testing (middle), the CS+ odour excites specific γm KCs (thick grey arrow), which in turn activate γd axons in γ4–γ5 compartments (grey dashed lines to yellow). Reverse γd-mediated activation of γm KCs occurs with unisensory colour testing (right).

Voltage-imaging of the γd KC somata did not show odour activation after multisensory reward training (Extended Data Fig. 3f,g), suggesting that learning-driven recruitment of these neurons to be odour-responsive does not occur by enhancing dendritic input. We therefore further tested a model of axonal recruitment. We reasoned that if DANs direct the recruitment of γd axons to become odour activated (and γm axons to become colour activated), an aversive learning event that requires dopamine release from the PPL1-γ1pedc DANs that innervate the most proximal γ1 lobe compartment[31] should confer odour responsiveness onto all downstream γd axon segments from γ1 to γ5. We confirmed that multisensory colour and odour aversive (electric shock) training produced memory enhancement for both combined and individual cues, similar to that observed after appetitive training (Extended Data Fig. 4a–f), and that γd KCs were also required for

odour memory enhancement following aversive multisensory training (Extended Data Fig. 4g–l). We next used two-photon voltage imaging to test whether γd axons from γ1 onwards gained CS+ odour activation after multisensory aversive learning. CS− odour presentation evoked hyperpolarization of γd KCs in both γ1 and γ5 compartments, regardless of training protocol (Fig. 4a,b and Extended Data Fig. 4m–p, light purple trace). By contrast, after multisensory aversive training, CS+ odour produced brief excitation of γd KCs in γ1 and γ5 (Fig. 4a,b, dark purple trace). Recordings of odour responses in γd KCs following appetitive and aversive multisensory learning, and of colour-evoked γm KCs after appetitive multisensory learning, are therefore consistent with the location of DAN teaching signals determining the portion of a γ-KC axon that gains activation by the other CS+ modality. Whereas aversive learning makes all axon segments downstream of γ1 excitable by the reciprocal modality (Fig. 4c), reward learning mostly alters CS+ excitation within the γ4 and γ5 segments (Fig. 3g).

Fig. 4: γd KCs become odour activated after aversive multisensory learning.



a,**b**, Aversive multisensory (colour + odour) training and odour imaging timelines (top left). The imaging plane in either the γ1 (**a**) or γ5 (**b**) region of γd KC axons (bottom left). Traces of CS+ and CS− odour-evoked activity (middle). Quantification of odour-evoked responses (right). The γ1 region showed excitation to the CS+ odour and inhibition to the CS− odour after aversive multisensory training (**a**). The CS+ odour evoked less inhibition in

γ1 than CS− (**b**). Odour-evoked activity traces show mean (solid line) with s.e.m. (shadow). Horizontal dashed lines indicate baseline activity. The solid black line below traces marks 5-s odour exposure. Asterisks denote significant difference between averaged CS+ and CS− responses ($P < 0.05$). CS+ and CS− responses for each fly are connected by a dashed line. **c**, MB model for aversive multisensory training followed by odour testing. Aversive multisensory training (left) engages punishment DANs (red) that depress synapses[16,17] between γm and γd KCs and approach-directing γ1 and γ2 MB output neurons[16,17] (not shown) while also binding γd and γm KC activity in these compartments. Unisensory odour testing (right) excites specific γm KCs, which activate γd axons from γ1 forward. **d**,**e**, Previous aversive multisensory learning enhances future appetitive but not aversive odour learning. Protocols are shown (left). Starved flies were divided into aversive multisensory training (group I) and aversive unisensory odour training (group II). Three hours later, both groups were trained with odours and sugar reward (using the same CS+ or CS− odours as for the initial training) and tested immediately afterwards. Memory performance is also shown (right). Group I initially trained with the multisensory aversive protocol performed better than group II initially trained with only odours (**d**). Group I initially trained with the multisensory appetitive protocol did not outperform group II initially trained with only odours (**e**). Asterisks denote significant differences ($P < 0.05$). Data are presented as mean ± s.e.m. Individual data points displayed as dots correspond to independent experiments. Groups were compared using paired two-sided *t*-test (**a**,**b**) and unpaired two-sided *t*-test (**d**,**e**); exact *P* values and comparisons are provided in Supplementary Information. $n = 24$ flies (**a**,**b**), $n = 12$ (**d**) and $n = 8$ (**e**).

Engram expansion benefits new learning

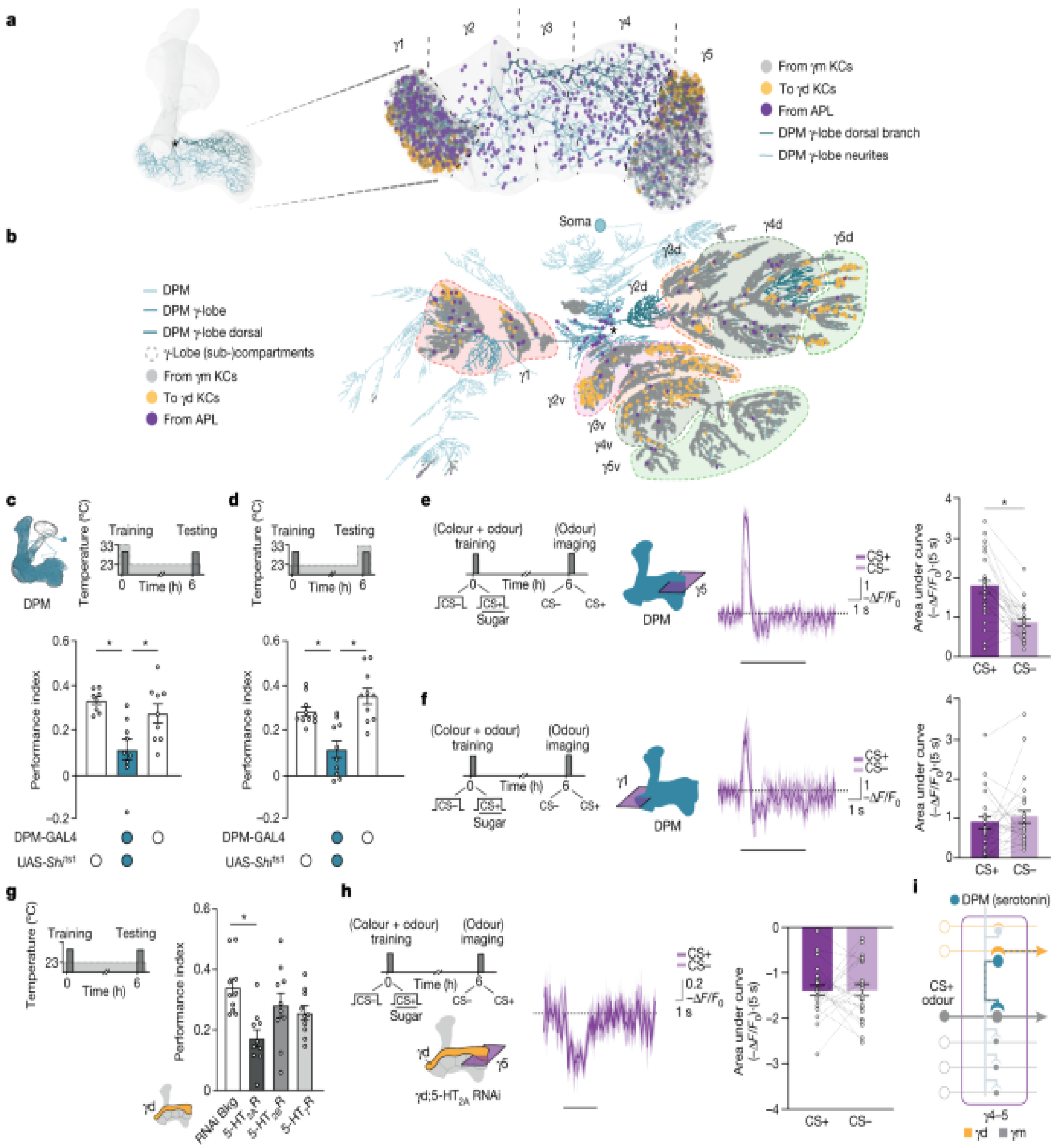
An expansion of the CS+ odour representation into a particular segment of the γd axons after multisensory training might be expected to facilitate subsequent learning with the same odour, if the next DAN teaching signal intersects the expanded KC representation. We tested this notion by sequentially training flies with either an aversive (dopamine in γ1) or appetitive (dopamine in γ4 and γ5) multisensory protocol followed by unisensory odour-reward or odour-punishment learning (Fig 4f,g). Previous multisensory aversive training significantly enhanced subsequent odour-

reward learning (Fig. 4d). However, no enhancement was apparent if aversive odour learning followed multisensory appetitive learning (Fig. 4e). Therefore, the multisensory training-dependent expansion of the CS+ odour representation can be included into the next CS+ odour memory engram if appropriate γd axon segments have become CS+ odour activated.

DPM neurons bridge KC sensory streams

The anatomy of the MB network suggests two possible ways to confer odour responsiveness to γd KC axons: via KC–KC synapses or neurons positioned to bridge the different KC streams. We queried the anatomical feasibility of these routes using the complete MB connectome of a single adult female fly 'hemibrain' electron microscope volume[32,33]. Although most (562 of 590) γm KCs make synapses with γd KCs, the number and placement of these connections do not support every γd KC to receive γm input in every γ-lobe compartment. In addition, KC–KC connections were reported to suppress activity in neighbouring KCs[34]. We next studied the fine anatomy of the γ-lobe innervation of the potentially excitatory serotonergic dorsal paired medial (DPM) neuron in the hemibrain electron microscope volume. DPM neurons send separate branches that densely innervate the vertical and horizontal lobes and distal peduncle of the MB, where they are both presynaptic and postsynaptic to KCs[35,36,37]. The ultrastructure of DPM neuronal projections in the γ-lobe revealed two branches within γ1 and other ventral and dorsal branches passing through the γ2–γ5 compartments (Fig. 5a). The positions of DPM neuronal synapses onto γd KCs follow the γd KC axon bundle as it winds around the γ-lobe from ventral in the γ1 compartment to dorsal in the γ5 compartment (Fig. 5a).

Fig. 5: DPM mediates multisensory stimulus binding.



a, 3D representation of MB (light grey) with DPM neuron γ-lobe neurites (teal) (left). DPM trifurcates (asterisk) into dorsal (dark teal), ventral (light teal) and γ1 compartment (teal) branches. Neurites are shaded by Strahler order, and twigs with Strahler order less than 1 were pruned. Details of the γ-lobe with γ1–γ5 compartment borders (dashed lines) are displayed (right). DPM presynapses to γd KCs (yellow spheres) colocalize on the dorsal DPM branch in γ5. Input synapses from γm KCs (grey spheres) to DPM are located throughout the ventral and dorsal branches. In γ5, 450 of 585 γm KCs made synapses with the DPM dorsal branch, where 89 of 98 γd KCs

also received DPM input. APL neuronal inputs (magenta spheres) localize along both DPM branches. **b**, 2D dendrogram projection of DPM neurites (shades of teal; see Extended Data Fig. 5a for details). The γ1 compartment is marked and γ2–γ5 compartments are split between dorsal and ventral DPM neuronal branches. Inputs from γm KCs (grey spheres) and outputs to γd KCs (yellow spheres) colocalize on compartment-specific branches. Inhibitory inputs from the APL neuron (magenta spheres) are distributed across DPM neurites. APL connectivity is detailed in Extended Data Fig. 5b. **c**,**d**, DPM neuron schematic (**c**; top). The timeline with temperature shifting (dashed line) is also shown. Blocking output of DPM neurons with VT64246-GAL4;UAS-*Shi*^{ts1} during training (**c**) or testing (**d**) of the 6-h olfactory retrieval performance is displayed (bottom). **e**,**f**, Appetitive multisensory (colour + odour) training and odour imaging timelines for DPM neurons (left). The imaging planes in γ5 (**e**) and γ1 (**f**) regions of DPM neurons and traces of CS+ and CS− odour-evoked activity are shown (middle). Quantification of responses is displayed (right). γ5 and γ1 regions of DPM showed excitatory responses to CS+ and CS− odours, but CS+ responses were only specifically enhanced in γ5. **g**, Timeline (left). RNAi knockdown of 5-HT$_{2A}$, but not 5-HT$_{2B}$ or 5-HT$_7$, receptors in γd KCs with MB607B-GAL4 impaired olfactory retrieval performance after multisensory training. In **c**,**d**,**g**, data are represented as mean ± s.e.m., individual data points are displayed as dots and the asterisks denote significant difference ($P < 0.05$). Bkg, RNAi background. **h**, Appetitive multisensory training and odour imaging timeline (top left). The imaging plane in γ5 of γd KC axons (bottom left). Traces of CS+ and CS− odour-evoked activity (middle). 5HT$_{2A}$R-RNAi eliminated gain-of-odour-evoked excitation in γ5 of γd KCs after multisensory training (see also Fig. 3a); CS+ and CS− odours similarly inhibited γd KC axons. Quantification is also shown (right). Odour-evoked activity traces show mean (solid line) with s.e.m. (shadow). Horizontal dashed lines indicate baseline activity. The solid black line underneath traces marks 5s odour exposure. The asterisk in **e** denotes significant difference between averaged CS+ and CS− responses ($P < 0.05$). CS+ and CS− responses for each fly are connected by dashed lines. Groups were compared using one-way ANOVA with Tukey's test (**c**,**d**), paired two-sided *t*-test (**e**,**f**,**h**) and one-way ANOVA with Dunnett's test (**g**); exact *P* values and comparisons are provided in Supplementary Information. $n = 8$ for *Shi* and $n = 9$ for other groups (**c**); $n = 10$ (**d**,**g**); $n = 26$ flies (**e**,**f**); and $n = 24$ flies (**h**).

See Extended Data Fig. 7 for controls. **i**, Model of DPM microcircuit bridging of odour-specific and colour-specific KCs following multisensory learning.

Annotating a dendrogram of DPM neurites (Fig. 5b) with γ-lobe compartment boundaries (based on DAN connectivity), synapses from γm KCs and those to γd KCs, showed that unique branches of the DPM neuron can provide compartment-specific microcircuit bridges between γm and γd KCs. DPM neurons can also bridge γd to γm KCs (Extended Data Fig. 5a). The large GABAergic anterior paired lateral (APL)[38] neuron was found to make synapses along DPM branches in the γ-lobe (Fig. 5b and Extended Data Fig. 5a), suggesting that DPM bridging can be regulated by local inhibition. The APL neuron receives many DAN inputs within each compartment and can therefore also be regulated with region specificity[39] (Extended Data Fig. 5b), to potentially release specific DPM branches from APL inhibition.

We challenged this putative microcircuit bridge model by independently manipulating APL and DPM neurons. Expression in APL neurons of the DopR2 dopamine receptor has been linked to aversive learning[40], and transcriptional profiling has suggested that APL neurons also express the DopEcR receptor[41]. Both of these dopamine receptors are known to have inhibitory action[42,43]. We therefore used *tub*P-GAL80^{ts} (ref. 44) to temporally restrict transgenic RNAi in APL neurons to test a role of these receptors in multisensory learning. Knocking down *Dop2R* in adult APL neurons abolished multisensory enhancement of olfactory retrieval performance (Extended Data Fig. 6a,c). A mild defect was also observed for odour memory following olfactory appetitive conditioning; however, the difference was only significant to one control (Extended Data Fig. 6b,d). By contrast, *DopEcR* RNAi had no effect in either experiment (Extended Data Fig. 6e,f). These results are consistent with reinforcing dopamine-inhibiting APL neurons to allow recruitment of γd KCs into the olfactory memory engram during multisensory learning.

We tested a role for the serotonergic DPM neurons using expression of UAS-*Shi*^{ts1}. Temporally restricting transmission from DPM neurons either during acquisition (Fig. 5c and Extended Data Fig. 7a) or retrieval (Fig. 5d

and Extended Data Fig. 7a) significantly impaired the multisensory training enhancement of odour retrieval memory. Blocking DPM neuronal output also impaired retrieval of visual memory after multisensory learning (Extended Data Fig. 7b). These same manipulations had no effect on odour memory after unisensory olfactory learning (Extended Data Fig. 7c), as in a previous study[45]. These data are consistent with DPM neuronal output being required during learning to bind together simultaneously active KC streams, whereas DPM neuronal output during memory retrieval provides the connection for odour-driven γm KCs to activate the relevant γd KCs.

We next performed behavioural and physiological experiments to directly test a model that multisensory learning establishes an excitatory DPM neuron microcircuit bridge between olfactory γm and visual γd KCs. We first used UAS-*Shi*^{ts1} to determine whether γ-KC (γm and γd) output was required during training for enhancement of olfactory and visual memory retrieval after multisensory training. Whereas blocking γ-KCs during multisensory learning significantly impaired both olfactory and visual retrieval (Extended Data Fig. 7e–h), it did not alter olfactory learning (Extended Data Fig. 7d), as in previous studies[46,47,48]. Finding that network plasticity of multisensory memory requires KC output suggests that it involves different learning rules to that of unisensory olfactory memory[16,49].

We used UAS-ASAP2f to search for compartment-specific plasticity of the functional connectivity of DPM neurons following multisensory reward learning. Olfactory reward learning was shown to specifically increase the calcium responses in DPM neurons to the CS+ odour for up to 2.5 h (refs. [36,50]) (see 1-h voltage recordings; Extended Data Fig. 7k–l). Recording 6 h after multisensory training revealed a clear and specific increase of CS+ odour voltage responses in DPM projections in the γ5, but not the γ1, compartment (Fig. 5e,f), which was absent at this time following olfactory learning (Extended Data Fig. 7i,j). This suggests that multisensory reward learning potentiates synapses from CS+ odour-specific γm KCs to DPM neurons within the microcircuitry of the γ5 compartment of the MB. As our previous imaging of γd KCs showed that they too become CS+ odour activated in their γ5 segments after multisensory reward learning (Fig. 3a), we tested whether gain-of-γd odour responsiveness could be mediated by DPM neuron-released serotonin (5-hydroxytryptamine (5-HT)). We first

established that bath application of 5-HT (in the presence of tetrodoxin to block indirect activation via other neurons) directly evoked depolarization of γd KCs expressing UAS-ASAP2f in naive flies (Extended Data Fig. 7o,p). 5-HT can exert excitatory effects through 5-HT_{2A}-type and 5-HT_{7}-type receptors[51]. We therefore used RNAi to knockdown these receptors in γd KCs. Reducing 5-HT_{2A}, but not 5-HT_{7} or 5-HT_{2B}, receptor expression impaired olfactory memory performance after multisensory training (Fig. 5g) but not olfactory training (Extended Data Fig. 7q). In addition, co-expressing *5-HT2A* RNAi with UAS-ASAP2f in γd KCs abolished the multisensory learning-induced gain-of-CS+ odour activation in the γ5 region of γd KCs (Fig. 5h), but did not affect colour-evoked responses (Extended Data Fig. 7r). Together, these anatomical, genetic and physiological data lead us to conclude that reinforcer-evoked compartment-specific dopamine releases APL-mediated inhibition, which facilitates the same reinforcing dopamine to induce KC–DPM and DPM–KC plasticity that forms excitatory serotonergic odour–colour-specific DPM microcircuit bridges between the relevant γm and γd KCs (Fig. 5i), and probably vice versa.

Discussion

Our study describes a precise neural mechanism in *Drosophila* through which multisensory learning improves subsequent memory performance, even for individual sensory cues. A single training trial with visual cues could only generate robust memory performance if they were combined with odours during training, similar to visual rhythm perception learning in humans, which requires accompanying auditory information[52]. We showed that multisensory learning binds together information from temporally contingent odours and colours within axons of MB γ-KCs, via serotonergic DPM neurons, whose activity also defines the coincidence time window[53]. This learning-driven binding converts axons of visually (presumably colour) selective KCs to also become responsive to the temporally contingent trained odour. We also demonstrated that axons of olfactory-selective KCs become activated by the temporally contingent trained colour. Although predominant dendritic input defines γm KCs as being olfactory and γd as being visual, our recordings showed that segments of their axons become multimodal after multisensory learning. This result suggests that γ-KCs are a

likely substrate where other temporally contingent sensory information can be integrated with that of explicit sensory cues[54,55].

Although our experiments mostly focused on odour-activated γm KCs recruiting colour γd KCs via DPM microcircuits, the observed behavioural enhancement of visual memory following multisensory learning, the demonstration that γm KCs become responsive to the trained colour, and the reciprocal connectivity of DPM neurons suggest that DPM neurons also probably mediate a reverse polarity bridge. In so doing, multisensory learning uses DPM neurons to link KCs that are responsive to each temporally contingent sensory cue and expands representations of each cue into that of the other. This cross-modal expansion allows multisensory experience to be efficiently retrieved by combined cues and by each individually. As a result, trained flies can evoke a memory of a visual experience with the learned odour, and memory of an odour with the learned colour. These findings provide a neural mechanism through which the fly achieves a conceptual equivalent of hippocampus-dependent pattern completion in mammals, in which partial scenes can retrieve a more complete memory representation[56]. Human patients with schizophrenia and autism exhibit deficits in multisensory integration[57], and these conditions have been linked to serotonergic dysfunction and 5-HT_{2A} receptors[58]. Our work here suggests that inappropriate routing of multisensory percepts may contribute to these conditions. Moreover, the excitatory 5-HT_{2A} receptors that mediate multisensory binding are the major targets of hallucinogenic drugs[59].

Methods

Fly strains

All *Drosophila melanogaster* strains were reared at 25 °C and 40–50% humidity, except where noted, on standard cornmeal-agar food (100 g l^{-1} anhydrous d-glucose, 47.27 g l^{-1} organic maize flour, 25 g l^{-1} autolysed yeast, 7.18 g l^{-1} agar and 12.18 g Tegosept dissolved in 8.36 ml absolute ethanol, per litre of fly food) in 12:12-h light:dark cycle. Canton-S flies were used as wild type (WT) and originated from William Quinn's laboratory

(Massachusetts Institute of Technology, Cambridge, MA, USA). The following GAL4 lines were used in the behavioural experiments: MB607B-GAL4 (refs. [13,60]), MB009B-GAL4 (refs. [13,60]), c708a-GAL4 (ref. [61]), VT43924-GAL4.2 (ref. [39]) and VT64246-GAL4 (ref. [62]). Temperature-controlled blocking of neuronal output was achieved by expressing the UAS-*Shi*^{ts1} (ref. [21]) transgene under the control of the MB607B-GAL4 (refs. [13,60]), MB009B-GAL4 (refs. [13,60]), c708a-GAL4 (ref. [61]) and VT64246-GAL4 (ref. [62]) drivers. For RNAi knockdown experiments involving APL, *tub*P-GAL80^{ts} (ref. [44]), VT43924-GAL4.2 (ref. [39]) flies were crossed with UAS-Dop2R RNAi[63] and UAS-DopEcR RNAi (VDRC ID: 103494) flies. The same driver line was crossed with WT flies and the RNAi background strain (VDRC ID: 60100), as controls. For RNAi knockdown experiments involving γd KCs, MB607B-GAL4 (refs. [13,60]) flies were crossed with UAS-5-HT_{2A} RNAi (31882, BDSC), UAS-5-HT_{2B} RNAi (60488, BDSC) and UAS-5-HT_7 RNAi (27273, BDSC) flies. The same driver line was crossed with the RNAi background strain (36304, BDSC), as controls. For live-imaging experiments, UAS-ASAP2f[64] was expressed using MB607B-GAL4 (refs. [13,60]) and VT64246-GAL4 (ref. [62]) and UAS-ASAP2s[65] with the 1471-GAL4 (ref. [66]) driver line. We used both male and female flies for the behavioural and imaging experiments.

Behavioural experiments

Male flies from the GAL4 lines were crossed to UAS-*Shi*^{ts1} virgin females, except for experiments involving c708a-GAL4, in which UAS-*Shi*^{ts1} males were crossed with c708a-GAL4 virgin females. For heterozygous controls, GAL4 or UAS-*Shi*^{ts1} flies were crossed to WT flies. In RNAi experiments, GAL4 or RNAi flies were crossed with the appropriate RNAi background strains or WT flies, respectively. All flies were raised at 25 °C, except where noted below for manipulation of RNAi expression. Populations of 2–8-day-old flies were used in all experiments.

For appetitive conditioning experiments, 80–100 flies were placed in a 25-ml vial containing 1% agar (as a water source) and a 20 × 60-mm piece of filter paper for 19–22 h before training and were kept starved for the entire experiment, except when assaying 24-h memory in which flies were fed for

30 min after training then returned to starvation vials until testing. For aversive conditioning experiments, 80–100 flies were placed in a vial containing standard food and a piece of filter paper for 14–22 h before behavioural experiments.

For experiments involving neuronal blocking with UAS-*Shi*^{ts1}, a schematic of the timeline of temperature shifting is provided in each figure. For *Shi*^{ts1} experiments, flies were transferred to a restrictive 33 °C for 30 min before training and/or testing. For RNAi experiments involving *tub*P-GAL80^{ts};VT43924-GAL4.2, flies were raised at 18 °C and shifted to 29 °C after eclosion to induce RNAi expression for 3 days before the behavioural experiments. The flies remained at 29 °C for the duration of the experiments.

All behavioural experiments were conducted using a standard T-maze that was modified to allow simultaneous delivery of colour and odour stimuli. The T-maze, which is made from translucent plastic, was covered in opaque blackout film to minimize interference between the visual stimuli when they were used in parallel. Odours were MCH and OCT diluted in mineral oil (at approximately 1:10^{-3} dilution). Colours were provided by light-emitting diodes (LEDs); green LEDs with a wavelength of 530 ± 10 nm (PM2E-3LGE-SD, ProLight Opto) and blue LEDs with a wavelength of 465 ± 10 nm (PM2B-3LDE-SD, ProLight Opto). Four LEDs were assembled in a circuit built onto a heat sink and were mounted securely on top of the odour delivery tubes. The intensities of the LEDs were adjusted so that naive flies showed no phototactic preference between the illuminated T-maze arms. Visual stimuli were presented in the same manner and same intensity for both training and testing. For appetitive experiments, the testing tubes were lined with filter paper; for aversive experiments, the testing tubes were lined with non-electrified shock grids. Experiments were performed in an environmental chamber set to the desired temperature and 55–65% relative humidity. Flies were handled before training and testing under overhead red light.

Appetitive conditioning was performed essentially as previously described[67]. In brief, flies were exposed for 2 min to stimuli Y (Y_{Colour} and/or Y_{Odour}) without reinforcement in a tube with dry filter paper (CS−), 30 s of clean air, then 2 min with stimuli X (X_{Colour} and/or X_{Odour}) presented with

5.8 M sucrose dried on filter paper (CS+). For aversive olfactory conditioning[17,18], flies received 1-min exposure to stimuli X (X_{Colour} and/or X_{Odour}) paired with twelve 90-V electric shocks at 5-s intervals (CS+), 45 s of clean air, followed by 1-min exposure to stimuli Y (Y_{Colour} and/or Y_{Odour}) without reinforcement (CS−). Electric shocks were delivered using a Grass S48 Square Pulse Stimulator (Grass Technology). Shock grids were those previously described[68] and consist of interleaved copper rows printed on transparent Mylar film, which allows coloured light to pass through.

Memory performance was assessed by testing flies for their preference between the CS− and the CS+ colours and/or odours for 2 min. Odour testing was performed in darkness. The flies in each arm were collected and transferred to polystyrene tubes (14-ml round bottom polypropylene test tube with cap, Falcon). Tubes with flies were frozen at −20 °C and flies were then removed and manually counted.

Performance indices were calculated as the number of flies in the CS+ arm minus the number in the CS− arm, divided by the total number of flies. For all behavioural experiments, a single sample, or *n*, represents the average performance index from two independent groups of flies trained with the reciprocal colour–odour combinations as CS+ and CS−. The total *n* for each experiment was acquired over three different training sessions on different days.

Six behavioural protocols were used:

1. (1)

   Visual learning: colours (X_{Colour} and Y_{Colour}) were used as CS+ and CS−.

2. (2)

   Olfactory learning: odours (X_{Odour} and Y_{Odour}) were used as CS+ and CS−.

3. (3)

Congruent protocol: colours and odours were combined (X_{Colour} + X_{Odour} and Y_{Colour} + Y_{Odour}) as CS+ and CS−. The same colour and odour combinations were used during training and testing.

4. (4)

   Incongruent protocol: colour and odour stimulus contingencies were switched between training (X_{Colour} + Y_{Odour} and Y_{Colour} + X_{Odour}) and testing (X_{Colour} + X_{Odour} and Y_{Colour} + Y_{Odour}).

   The visual and olfactory learning protocols are unisensory, whereas the congruent and incongruent protocols are multisensory.

5. (5)

   Olfactory retrieval: flies were trained as in the congruent protocol, but only odours (X_{Odour} and Y_{Odour}) were presented as the choice at test.

6. (6)

   Visual retrieval: flies were trained as in the congruent protocol, but only colours (X_{Colour} and Y_{Colour}) were presented as the choice at test.

The sequential learning experiments depicted in Fig. 4f,g used aversive or appetitive congruent multisensory training followed by unisensory appetitive or aversive olfactory learning, then testing using olfactory retrieval.

Two-photon voltage imaging

All flies were raised at 25 °C and 3–8-day-old male and female flies were used in all experiments. Imaging experiments were performed essentially as previously described[69,70,71]. In brief, flies were trained in the T-maze setup using either olfactory learning (protocol 2), a congruent multisensory protocol (protocol 3) or an unpaired training protocol. In unpaired training, flies were exposed to the combined odour and visual stimuli (X_{Colour} + X_{Odour} and Y_{Colour} + Y_{Odour} combination), but the shock or sugar was presented alone 2 min before or after the CS+, respectively. After training,

flies were kept in darkness until recording. Just before recording, flies were briefly immobilized on ice and mounted in a custom-made chamber allowing free movement of the antennae and legs. The head capsule was opened under room temperature carbogenated (95% O_2 and 5% CO_2) buffer solution, and the fly, in the recording chamber, was placed under a two-photon microscope (Scientifica). For starved flies, the following sugar-free buffer was used: 108 mM NaCl, 5 mM KCl, 5 mM HEPES, 15 mM ribose, 4 mM $NaHCO_3$, 1 mM NaH_2PO_4, 2 mM $CaCl_2$ and 8.2 mM $MgCl_2$, osmolarity 272 mOsm, pH 7.3). For fed flies, the following buffer was used: 103 mM NaCl, 3 mM KCl, 5mM N-Tris, 10 mM trehalose, 10 mM glucose, 7 mM sucrose, 26 mM $NaHCO_3$, 1 mM NaH_2PO_4, 1.5 mM $CaCl_2$ and 4 mM $MgCl_2$, osmolarity 275 mOsm, pH 7.3).

Flies were subjected to a constant air stream, carrying vapour from mineral oil solvent (air). For odour-evoked imaging experiments, flies were sequentially exposed to CS+ and CS− odour, each for 5 s, interspersed by 30 s, to simulate the behavioural test. As in the behaviour experiments, the odours were MCH and OCT (diluted in mineral oil at approximately $1:10^{-3}$), and they were used reciprocally as CS+ and CS−. Any flies that did not respond to one of the two presented odours were excluded from further analyses. For colour-evoked imaging experiments, the colour presentation was interleaved with image acquisition. This was achieved using a shutter on the objective (Ø1/2" stainless steel diaphragm optical beam shutter with controller, Thorlabs) and a second externally controlled shutter (Vincent/UniBlitz VS35S2ZM1R1-21 Uni-Stable Shutter; UniBlitz VMM-T1 Shutter Driver/Timer Controller) on the LED delivery system. For each cycle of recording, colour was presented for 0.75 s at 0.4 Hz and followed by image acquisition for 1.75 s. Importantly, 0.4-Hz pulsed colour presentation evoked robust responses in γd KCs in naive flies, measured with UAS-ASAP2f (Extended Data Fig. 4h), and behavioural memory testing with 0.4-Hz flickered colours produced similar memory performance to that generated with continuous colour presentation (Extended Data Fig. 4d,e). We used UAS-ASAP2s for γm KC recordings because it produces slower and larger responses than ASAP2f, which we considered to be beneficial for the image acquisition being interleaved with colour presentation. Flies were sequentially exposed four times to the CS+ colour and then four times to the CS− colour with each colour presentation followed by an image acquisition

cycle. A 30-s interval separated the CS+ and CS− recordings. Blue and green were used reciprocally as CS+ and CS−. One hemisphere of the brain was randomly selected to image KC axons. It is rarely possible to image across all MB compartments of the γ-lobe because γ1 and γ5 are most often in different planes. We therefore had to analyse these two compartments independently.

Fluorescence was excited using approximately 140-fs pulses, 80-MHz repetition rate, centred on 910 nm generated by a Ti-Sapphire laser (Chameleon Ultra II, Coherent). Images of 256 × 256 pixels were acquired at 5.92 Hz, controlled by ScanImage 3.8 software[72]. Odours were delivered using a custom-designed system[73], controlled by LabView (v.11).

For acute 5-HT application, we used a perfusion pump system (14-284-201, Fisher Scientific) to continuously deliver saline at a rate of approximately 0.043 ml s^{-1}. 5-HT was applied in the presence of 1 μM tetrodotoxin to block voltage-gated sodium channels and propagation of action potentials that could result in indirect excitation. To examine the effects of serotonin on γd KC membrane voltage, baseline fluorescence was recorded for 5 min before switching to a solution containing 100 μM serotonin hydrochloride (H9523, Sigma Aldrich) for an additional 5 min of recording. Washout was performed by changing the solution back to saline. The time of application and concentration of 5-HT used is comparable to recent physiological studies applying exogenous 5-HT to the *Drosophila* brain[74,75,76,77]. Owing to perfusion tubing length and dead volume, the perfusion switch took approximately 70 s to reach the brain.

For analysis, two-photon fluorescence images were manually segmented using Fiji[78], using a custom-made code including an image stabilizer plugin[79]. Movement of the animals was small enough for images to not require registration. For subsequent quantitative analyses, custom Fiji and MATLAB scripts were used. The baseline fluorescence, F_0, was defined for each stimulus response as the mean fluorescence F from 2 s before and up to the point of odour or colour presentation (or 30 s after the start of the recordings for 5-HT treatments). $-\Delta F/F_0$ accordingly describes the fluorescence relative to this baseline. For the odour-evoked responses of KCs, the area under the curve was measured as the integral of $-\Delta F/F_0$

during the 5-s odour stimulation. We elected to maintain the natural units of the experiment when reporting the integrated area under the curve (that is, $(-\Delta F/F_0) \times (5\ \mathrm{s})$), because we do not make any inferences regarding the shape of the response. For the colour-evoked KC responses, the mean fluorescence signal ($-\Delta F/F_0$) for the first acquisition cycle (1.75 s) was quantified. Each *n* corresponds to a recording from a different individual fly. All data were acquired over three different training sessions on different days.

For 5-HT treatments, we defined the 'pre-treatment' as the mean $-\Delta F/F_0$ value for 300 s before the 5-HT delivery, the 5-HT application was the mean $-\Delta F/F_0$ for 300 s during 5-HT delivery and the 'washout' treatment as the mean $-\Delta F/F_0$ for 300 s from the offset of drug delivery. Traces were smoothed over 5 s by a moving average filter. Each *n* corresponds to a recording from a different individual fly. All data were acquired across three different imaging sessions on different days.

ASAP2f and ASAP2s data are presented as $-\Delta F/F_0$ to correct the inverse relation between sensor fluorescence and membrane voltage.

Statistical analysis

Statistical analyses were performed in GraphPad Prism. All behavioural data were analysed with an unpaired two-sided *t*-test, Mann–Whitney *U*-test, one-way ANOVA or Kruskal–Wallis *H*-test followed by a post-hoc Tukey's, Dunnett's or Dunn's multiple comparisons test. No statistical methods were used to predetermine sample size. Sample sizes were similar to other publications in the field. For the imaging experiments, odour-evoked responses were compared by a paired two-sided *t*-test for normally distributed data and repeated measures one-way ANOVA, and Wilcoxon signed-rank test was used for non-Gaussian distributed data. Normality was tested using the D'Agostino and Pearson normality test. For imaging data, a method for outlier identification was run for each dataset (ROUT method), which is based on the false discovery rate. The false discovery rate was set to the highest *Q* value possible (10%). In datasets in which potential outliers were identified, statistical analyses were performed by removing all odour-evoked responses for those flies. The analyses with or without the outliers

were not different, so we decided to maintain and present the complete datasets, which may contain potential outliers. Partial eta squared was used to report effect sizes ($\eta^2 = 0.01$ indicates a small effect; $\eta^2 = 0.06$ indicates a medium effect; $\eta^2 = 0.14$ indicates a large effect); the formula used is reported in statistics table. All statistical analyses are also reported in the statistics table in the Supplementary Information.

Blinding and randomization

The experiments were randomized with appropriate controls present in each independent experiment. All genotypes tested and analysed were self-blinded to the experimenter. More details regarding research design are in the Reporting Summary.

Neuroanatomy, connectivity and dendrograms

Neuromorphological calculations and connectivity analyses were performed, and dendrograms were calculated and plotted, with scripts based on NAVis 1.2.1 library functions in Python 3.8.8 (https://pypi.org/project/navis/; https://github.com/navis-org/navis)[80] and data from the *Drosophila* hemibrain (v.1.2.1) (https://neuprint.janelia.org)[32,33]. All neuronal skeletons were healed (navis.heal_skeleton (method = "ALL", max_dist = "100 nanometer", min_size = 10)), rerooted (navis.reroot_skeleton (x.soma)) and strongly down sampled with conserved connectors (navis.downsample_neuron(downsampling_factor = 1000, preserve_nodes = 'connectors')).

3D representations of neurons shaded by Strahler order were generated with navis.plot2d (method='3d', shade_by='strahler_index'), after pruning twigs with Strahler order of 1 or less (navis.prune_by_strahler()). Where applicable, only branches in specific volumes were considered (navis.in_volume()). Volumes were obtained from neuprint (v.1.2.1) with fetch_roi(). Connectivity was analysed using unpruned neurons and with compartment specificity (navis.in_volume()).

Custom scripts based on navis.plot_flat() were used to generate dendrograms of DPM and APL neurons with twigs of Strahler order of 1 or less pruned.

MB compartment boundaries were defined by connectivity to DANs of the respective compartments. Branches outside the γ-lobe were downsized manually to increase the visibility of γ-lobe compartments. Synapses are filtered by in_volume() and displayed on branches with Strahler order of more than 1. Connectivity statistics are based on unpruned neurons, and synapses between neurons were obtained with R based natverse:: neuprint_get_synapses() (https://natverse.org)[80] scripts and processed with custom scripts in Python.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Data supporting the findings of this study are available at https://github.com/PJZO/Multisensory-learning-Engram.git. The dataset for the *Drosophila* hemibrain mentioned in the section 'Neuroanatomy, connectivity and dendrograms' in the Methods is publicly available at https://neuprint.janelia.org.

Code availability

Customized MATLAB and Fiji/ImageJ scripts are available at https://github.com/PJZO/Multisensory-learning-Engram.git. The code mentioned in the section 'Neuroanatomy, connectivity and dendrograms' in the Methods is publicly available at https://pypi.org/project/navis/, https://github.com/navis-org/navis and https://natverse.org.

References

1. Stein, B. E., Stanford, T. R. & Rowland, B. A. Development of multisensory integration from the perspective of the individual neuron. *Nat. Rev. Neurosci.* **15**, 520–535 (2014).

2. Guo, J. & Guo, A. Crossmodal interactions between olfactory and visual learning. *Science* **309**, 307–310 (2005).

3. Shams, L. & Seitz, A. R. Benefits of multisensory learning. *Trends Cogn. Sci.* **12**, 411–417 (2008).

4. Fetsch, C. R., Deangelis, G. C. & Angelaki, D. E. Bridging the gap between theories of sensory cue integration and the physiology of multisensory neurons. *Nat. Rev. Neurosci.* **14**, 429–442 (2013).

5. Thiagarajan, D. et al. Aversive bimodal associations differently impact visual and olfactory memory performance in *Drosophila*. *iScience* **25**, 105485 (2022).

6. Ghazanfar, A. A. & Schroeder, C. E. Is neocortex essentially multisensory? *Trends Cogn. Sci.* **10**, 278–285 (2006).

7. Driver, J. & Noesselt, T. Multisensory interplay reveals crossmodal influences on 'sensory-specific' brain regions, neural responses, and judgments. *Neuron* **57**, 11–23 (2008).

8. von Kriegstein, K. et al. Simulation of talking faces in the human brain improves auditory speech recognition. *Proc. Natl Acad. Sci. USA* **105**, 6747–6752 (2008).

9. Schall, S., Kiebel, S. J., Maess, B. & Von Kriegstein, K. Early auditory sensory processing of voices is facilitated by visual mechanisms. *Neuroimage* **77**, 237–245 (2013).

10. Vincis, R. & Fontanini, A. Associative learning changes cross-modal representations in the gustatory cortex. *eLife* **5**, e16420 (2016).

11. Knöpfel, T. et al. Audio–visual experience strengthens multisensory assemblies in adult mouse visual cortex. *Nat. Commun.* **10**, 5684 (2019).

12. Han, X., Xu, J., Chang, S., Keniston, L. & Yu, L. Multisensory-guided associative learning enhances multisensory representation in primary auditory cortex. *Cereb. Cortex* **32**, 1040–1054 (2022).

13. Aso, Y. et al. The neuronal architecture of the mushroom body provides a logic for associative learning. *eLife* **3**, e04577 (2014).

14. Li, F. et al. The connectome of the adult *Drosophila* mushroom body provides insights into function. *eLife* **9**, e62576 (2020).

15. Owald, D. et al. Activity of defined mushroom body output neurons underlies learned olfactory behavior in *Drosophila*. *Neuron* **86**, 417–427 (2015).

16. Hige, T. et al. Heterosynaptic plasticity underlies aversive olfactory learning in *Drosophila*. *Neuron* **88**, 985–998 (2015).

17. Perisse, E. et al. Aversive learning and appetitive motivation toggle feed-forward inhibition in the *Drosophila* mushroom body. *Neuron* **90**, 1086–1099 (2016).

18. Tully, T. & Quinn, W. G. Classical conditioning and retention in normal and mutant *Drosophila melanogaster*. *J. Comp. Physiol. A* **157**, 263–277 (1985).

19. Vogt, K. et al. Direct neural pathways convey distinct visual information to *Drosophila* mushroom bodies. *eLife* **5**, e14009 (2016).

20. Vogt, K. et al. Shared mushroom body circuits underlie visual and olfactory memories in *Drosophila*. *eLife* **3**, e02395 (2014).

21. Kitamoto, T. Conditional modification of behavior in *Drosophila* by targeted expression of a temperature-sensitive shibire allele in defined neurons. *J. Neurobiol.* **47**, 81–92 (2001).

22. Perisse, E. et al. Different Kenyon cell populations drive learned approach and avoidance in *Drosophila*. *Neuron* **79**, 945–956 (2013).

23. Yang, H. H. H. et al. Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell* **166**, 245–257 (2016).

24. Burke, C. J. et al. Layered reward signalling through octopamine and dopamine in *Drosophila*. *Nature* **492**, 433–437 (2012).

25. Liu, C. et al. A subset of dopamine neurons signals reward for odour memory in *Drosophila*. *Nature* **488**, 512–516 (2012).

26. Schwaerzel, M. et al. Dopamine and octopamine differentiate between aversive and appetitive olfactory memories in *Drosophila*. *J. Neurosci.* **23**, 10495–10502 (2003).

27. Claridge-Chang, A. et al. Writing memories with light-addressable reinforcement circuitry. *Cell* **139**, 405–415 (2009).

28. Aso, Y. et al. Three dopamine pathways induce aversive odor memories with different stability. *PLoS Genet.* **8**, e1002768 (2012).

29. Krashes, M. J. et al. A neural circuit mechanism integrating motivational state with memory expression in *Drosophila*. *Cell* **139**, 416–427 (2009).

30. Plaçais, P. Y. & Preat, T. To favor survival under food shortage, the brain disables costly memory. *Science* **339**, 440–442 (2013).

31. Waddell, S. Neural plasticity: dopamine tunes the mushroom body output network. *Curr. Biol.* **26**, R109–R112 (2016).

32. Scheffer, L. K. et al. A connectome and analysis of the adult *Drosophila* central brain. *eLife* **9**, e57443 (2020).

33. Plaza, S. M. et al. neuPrint: an open access tool for EM connectomics. *Front. Neuroinform.* **16**, 896292 (2022).

34. Manoim, J. E., Davidson, A. M., Weiss, S., Hige, T. & Parnas, M. Lateral axonal modulation is required for stimulus-specific olfactory conditioning in *Drosophila*. *Curr. Biol.* **32**, 4438–4450.e5 (2022).

35. Waddell, S., Armstrong, J. D., Kitamoto, T., Kaiser, K. & Quinn, W. G. The amnesiac gene product is expressed in two neurons in the *Drosophila* brain that are critical for memory. *Cell* **103**, 805–813 (2000).

36. Yu, D., Keene, A. C., Srivatsan, A., Waddell, S. & Davis, R. L. *Drosophila* DPM neurons form a delayed and branch-specific memory trace after olfactory classical conditioning. *Cell* **123**, 945–957 (2005).

37. Takemura, S.-y. et al. A connectome of a learning and memory center in the adult *Drosophila* brain. *eLife* **6**, e26975 (2017).

38. Liu, X. & Davis, R. L. The GABAergic anterior paired lateral neuron suppresses and is suppressed by olfactory learning. *Nat. Neurosci.* **12**, 53–59 (2009).

39. Amin, H., Apostolopoulou, A. A., Suárez-Grimalt, R., Vrontou, E. & Lin, A. C. Localized inhibition in the *Drosophila* mushroom body. *eLife* **9**, e56954 (2020).

40. Zhou, M. et al. Suppression of GABAergic neurons through D2-like receptor secures efficient conditioning in *Drosophila* aversive olfactory learning. *Proc. Natl Acad. Sci. USA* **116**, 5118–5125 (2019).

41. Aso, Y. et al. Nitric oxide acts as a cotransmitter in a subset of dopaminergic neurons to diversify memory dynamics. *eLife* **8**, e49257 (2019).

42. Hearn, M. G. et al. A *Drosophila* dopamine 2-like receptor: molecular characterization and identification of multiple alternatively spliced variants. *Proc. Natl Acad. Sci. USA* **99**, 14554–14559 (2002).

43. Lark, A., Kitamoto, T. & Martin, J. R. Modulation of neuronal activity in the *Drosophila* mushroom body by DopEcR, a unique dual receptor for ecdysone and dopamine. *Biochim. Biophys. Acta Mol. Cell Res.* **1864**, 1578–1588 (2017).

44. McGuire, S. E., Le, P. T., Osborn, A. J., Matsumoto, K. & Davis, R. L. Spatiotemporal rescue of memory dysfunction in *Drosophila*. *Science* **302**, 1765–1768 (2003).

45. Keene, A. C., Krashes, M. J., Leung, B., Bernard, J. A. & Waddell, S. *Drosophila* dorsal paired medial neurons provide a general mechanism for memory consolidation. *Curr. Biol.* **16**, 1524–1530 (2006).

46. Trannoy, S., Redt-Clouet, C., Dura, J. M. & Preat, T. Parallel processing of appetitive short- and long-term memories in *Drosophila*. *Curr. Biol.* **21**, 1647–1653 (2011).

47. McGuire, S. E., Le, P. T. & Davis, R. L. The role of *Drosophila* mushroom body signaling in olfactory memory. *Science* **293**, 1330–1333 (2001).

48. Dubnau, J., Grady, L., Kitamoto, T. & Tully, T. Disruption of neurotransmission in *Drosophila* mushroom body blocks retrieval but not acquisition of memory. *Nature* **411**, 476–480 (2001).

49. Handler, A. et al. Distinct dopamine receptor pathways underlie the temporal sensitivity of associative learning. *Cell* **178**, 60–75 (2019).

50. Cervantes-Sandoval, I. & Davis, R. L. Distinct traces for appetitive versus aversive olfactory memories in DPM neurons of *Drosophila*. *Curr. Biol.* **22**, 1247–1252 (2012).

51. Nichols, D. E. & Nichols, C. D. Serotonin receptors. *Chem. Rev.* **108**, 1614–1641 (2008).

52. Barakat, B., Seitz, A. R. & Shams, L. Visual rhythm perception improves through auditory but not visual training. *Curr. Biol.* **25**, R60–R61 (2015).

53. Zeng, J. et al. Local 5-HT signaling bi-directionally regulates the coincidence time window for associative learning. *Neuron* https://doi.org/10.1016/j.neuron.2022.12.034 (2023).

54. Keene, A. C. et al. Diverse odor-conditioned memories require uniquely timed dorsal paired medial neuron output. *Neuron* **44**, 521–533 (2004).

55. Jacob, P. F. et al. Prior experience conditionally inhibits the expression of new learning in *Drosophila*. *Curr. Biol.* **31**, 3490–3503.e3 (2021).

56. Hunsaker, M. R. & Kesner, R. P. The operation of pattern separation and pattern completion processes associated with different attributes or

domains of memory. *Neurosci. Biobehav. Rev.* **37**, 36–58 (2013).

57. Cascio, C. J., Simon, D. M., Bryant, L. K., DiCarlo, G. & Wallace, M. T. in *Multisensory Perception: From Laboratory to Clinic* (Elsevier, 2019).

58. Muller, C. L., Anacker, A. M. J. & Veenstra-VanderWeele, J. The serotonin system in autism spectrum disorder: from biomarker to animal models. *Neuroscience* **321**, 24–41 (2016).

59. López-Giménez, J. F. & González-Maeso, J. Hallucinogens and serotonin 5-HT2A receptor-mediated signaling pathways. *Curr. Top. Behav. Neurosci.* **36**, 45–73 (2018).

60. Aso, Y. et al. Mushroom body output neurons encode valence and guide memory-based action selection in *Drosophila*. *eLife* **3**, e04580 (2014).

61. Zhu, S. et al. Gradients of the *Drosophila* chinmo BTB-zinc finger protein govern neuronal temporal identity. *Cell* **127**, 409–422 (2006).

62. Lee, P.-T. T. et al. Serotonin-mushroom body circuit modulating the formation of anesthesia-resistant memory in *Drosophila*. *Proc. Natl Acad. Sci. USA* **108**, 13794–13799 (2011).

63. Draper, I., Kurshan, P. T., McBride, E., Jackson, F. R. & Kopin, A. S. Locomotor activity is regulated by D2-like receptors in *Drosophila*: an anatomic and functional analysis. *Dev. Neurobiol.* **67**, 378–393 (2007).

64. Yang, H. H. H. et al. Subcellular imaging of voltage and calcium signals reveals neural processing in vivo. *Cell* **166**, 245–257 (2016).

65. Chamberland, S. et al. Fast two-photon imaging of subcellular voltage dynamics in neuronal tissue with genetically encoded indicators. *eLife* **6**, e25690 (2017).

66. Isabel, G., Pascual, A. & Preat, T. Exclusive consolidated memory phases in *Drosophila*. *Science* **304**, 1024–1027 (2004).

67. Krashes, M. J. & Waddell, S. Rapid consolidation to a radish and protein synthesis-dependent long-term memory after single-session appetitive olfactory conditioning in *Drosophila*. *J. Neurosci.* **28**, 3103–3113 (2008).

68. Quinn, W. G., Harris, W. A. & Benzer, S. Conditioned behavior in *Drosophila melanogaster*. *Proc. Natl Acad. Sci. USA* **71**, 708–712 (1974).

69. Owald, D. et al. Activity of defined mushroom body output neurons underlies learned olfactory behavior in *Drosophila*. *Neuron* **86**, 417–427 (2015).

70. Felsenberg, J. et al. Integration of parallel opposing memories underlies memory extinction. *Cell* **175**, 709–722.e15 (2018).

71. Jacob, P. F. & Waddell, S. Spaced training forms complementary long-term memories of opposite valence in *Drosophila*. *Neuron* **106**, 977–991.e4 (2020).

72. Pologruto, T. A., Sabatini, B. L. & Svoboda, K. ScanImage: flexible software for operating laser scanning microscopes. *Biomed. Eng. Online* **2**, 13 (2003).

73. Shang, Y., Claridge-Chang, A., Sjulson, L., Pypaert, M. & Miesenböck, G. Excitatory local circuits and their implications for olfactory processing in the fly antennal lobe. *Cell* **128**, 601–612 (2007).

74. Zhang, X. & Gaudry, Q. Functional integration of a serotonergic neuron in the *Drosophila* antennal lobe. *eLife* **5**, e16836 (2016).

75. Liu, C. et al. A serotonin-modulated circuit controls sleep architecture to regulate cognitive function independent of total sleep in *Drosophila*. *Curr. Biol.* **29**, 3635–3646.e5 (2019).

76. Alekseyenko, O. V. et al. Serotonergic modulation of aggression in *Drosophila* involves GABAergic and cholinergic opposing pathways. *Curr. Biol.* **29**, 2145–2156.e5 (2019).

77. Sampson, M. M. et al. Serotonergic modulation of visual neurons in *Drosophila melanogaster*. *PLoS Genet.* **16**, e1009003 (2020).

78. Schindelin, J. et al. Fiji: an open-source platform for biological-image analysis. *Nat. Methods* **9**, 676–682 (2012).

79. Li, K. The image stabilizer plugin for ImageJ. http://www.cs.cmu.edu/~kangli/code/Image_Stabilizer.html (2008).

80. Bates, A. S. et al. The natverse, a versatile toolbox for combining and analysing neuroanatomical data. *eLife* **9**, e53350 (2020).

Acknowledgements

We thank R. Brain, K. delos Santos, R. Busby, M. Moreno Gasulla, J.-P. Moszynski and F. Woods for technical assistance; and G. Rubin, FlyLight, B. Dickson, A. Lin and the Vienna *Drosophila* Resource Center, and the Bloomington Stock Center for flies. Z.O. was funded by an EMBO long-term postdoctoral fellowship (ALTF 311-2017). P.V.-G. was funded by the Sloane Robinson/Clarendon Scholarship given by the University's Clarendon Fund scheme and Keble College and from the Consejo Nacional de Ciencia y Tecnología (CONACYT). S.W. was funded by a Wellcome Principal Research Fellowship (200846), a Wellcome Discovery Award (225192), an ERC Advanced Grant (789274) and Wellcome Collaborative Awards (203261 and 209235).

Author information

Author notes

1. Nils Otto

   Present address: Institute of Anatomy and Molecular Neurobiology, Westfälische Wilhelms-Universität Münster, Münster, Germany

2. These authors contributed equally: Zeynep Okray, Pedro F. Jacob

Authors and Affiliations

1. Centre for Neural Circuits & Behaviour, University of Oxford, Oxford, UK

   Zeynep Okray, Pedro F. Jacob, Ciara Stern, Kieran Desmond, Nils Otto, Clifford B. Talbot, Paola Vargas-Gutierrez & Scott Waddell

Contributions

Z.O., P.F.J., C.B.T. and S.W. designed the research. Z.O., P.F.J., C.S., K.D., N.O. and P.V.-G. performed the experiments. Z.O., P.F.J., C.S., K.D. and N.O. analysed the data. S.W. provided resources. S.W., Z.O., P.F.J. and N.O. wrote the manuscript. S.W. supervised the study. S.W. and Z.O. acquired funding.

Corresponding authors

Correspondence to Zeynep Okray or Scott Waddell.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Memory performance is most robust when color and odor combinations are consistent during acquisition and retrieval.

a. *Top left*, multisensory protocols. Green and blue squares represent colors, light and dark gray squares represent OCT and MCH odors. Visual (V) learning: colors used as CS+ and CS−. Olfactory (O) learning: odors used as CS+ and CS−. Congruent (C) protocol: colors+odors were combined as CS+ and CS−. Same color+odor combinations used during training and testing. Incongruent (I) protocol: colors+odors were combined as CS+ and CS− but combinations were switched between training and testing. *Bottom left*, training and testing timeline. *Right*, 24 h memory performance for V, O, C and I protocols. Combining colors+odors in the congruent (C) protocol enhanced 24 h performance, compared to that obtained with unisensory V or O learning. Incongruent (I) pairing of colors and odors abolished the multisensory enhancement of 24 h memory. **b**. *Left*, training and testing timeline. *Middle*, multisensory protocols. *Right*, immediate memory performance. Flies showed a significantly higher memory following the C than the I protocol. **c**. *Left*, training and testing timeline. *Middle*, multisensory protocols: C protocol as described above; Congruent protocol using the same color (C−sc) combined with different odors as CS+ and CS− during training and testing. *Right*, the C protocol using distinct color+odor combinations for CS+ vs CS− resulted in higher immediate memory performance than the C-sc protocol using the same color with both odors. **d**. *Left*, training and testing timeline. *Middle*, multisensory protocols: Olfactory (O) learning as described above; Multisensory Retrieval (MSR): odors were CS+ and CS− during training and these same odors were combined with different colors during testing. *Right*, immediate memory performance evoked by MSR was not significantly reduced to that following for Olfactory learning and retrieval. **e**. *Left*, training and testing timeline. *Middle*, multisensory protocols: Congruent (C) protocol as described above;

Odor Retrieval (OR): colors+odors were CS+ and CS− during training and only odors were used during testing. *Right*, flies trained with multisensory stimuli performed better if they were tested with congruent multisensory stimuli compared to only one modality (in this case odor). Asterisks denote significant differences ($P < 0.05$). Data presented as mean ± standard error of mean (SEM). Individual data points displayed as dots correspond to independent experiments. Groups compared using one-way ANOVA with Tukey's test (**a**), unpaired two-sided *t*-test (**b**, **c**, **e**) and unpaired two-sided Mann–Whitney *U*-test (**d**), exact *P* values and comparisons are given in Supplementary Information. N values for each experiment are: **a**, **b**, **e**, n = 10; **c**, n = 8; **d**, n = 10 for OR and n = 8 for MSR.

Extended Data Fig. 2 Constitutively blocking γd or $\alpha\beta_p$ KC output impairs the visual component of multisensory memories.

a. *Top left*, schematic of γd KCs. *Bottom left*, training and testing timeline with constant restrictive temperature (dashed line). *Right*, 6 h memory performance following Olfactory learning is unchanged when γd KCs are blocked through the experiment using MB607B-GAL4; UAS-*Shi*^{ts1} **b** and **c**. Blocking γd KCs throughout the experiment significantly impaired memory in the Congruent protocol (**b**). The release from the interference effect of the Incongruent protocol did not reach significance (**c**). **d**. *Top left*, schematic of αβp KCs. *Bottom left*, training and testing timeline with constant restrictive temperature (dashed line). *Right*, blocking αβp KC output throughout the experiment using c708a-GAL4; UAS-*Shi*^{ts1} did not impair 6 h Olfactory learning. **e** and **f**. Memory performance for Congruent (**e**) and Incongruent (**f**) protocols changed significantly when αβp KCs were blocked during the experiment. **g** and **h**. *Top*, training and testing timeline with temperature shifting (dashed line). Blocking γd (**g**) and αβp (**h**) KCs impaired memory retrieved with Visual cues. **i**. *Top left*, schematic of γd KCs. *Top right*, training and testing timeline with constant permissive temperature (dashed line). **i-l** Memory performance in MB607B-GAL4; UAS-*Shi*^{ts1} flies following Congruent (**i**), Incongruent (**j**), Olfactory Retrieval (**k**) and Visual Retrieval (**l**) protocols was not affected when training and testing was performed at 23 °C. **m**. *Top left*, schematic of αβp KCs. *Bottom left*, training and testing timeline with constant permissive temperature (dashed line). **m-o** Memory performance in c708a-GAL4; UAS-*Shi*^{ts1} flies after Congruent

(**m**), Incongruent (**n**), and Visual Retrieval (**o**) protocols was not affected when training and testing was at 23 °C. Asterisks denote significant differences (*P* < *0.05*). Data presented as mean ± standard error of mean (SEM). Individual data points displayed as dots correspond to independent experiments. Groups compared using one-way ANOVA with Tukey's test (**a-g, i-o**), and Kruskal–Wallis *H*-test with Dunn's test (**h**), exact *P* values and comparisons are given in Supplementary Information. N values for each experiment are: **a—c**, **g**, **h**, n = 12; **d**, n = 12 for c708a and n = 10 for all other groups, **f**, n = 10; **i-o**, n = 8.

Extended Data Fig. 3 Multisensory aversive learning enhances memory for the combined and individual odor and color cues.

a and **b**. *Top left*, unpaired appetitive training and odor imaging protocol. Color+odor presentation was not contingent with sugar. *Bottom left*, imaging plane in the γ5 region (**a**) and γ1 (**b**) of γd KCs. *Middle*, traces of CS+ and CS− odor-evoked activity. In both regions of γd KCs, unpaired multisensory appetitive training does not alter odor responses. *Right*, quantification of odor-evoked responses. **c**. Color-evoked (blue, green and control, i.e. LED off) responses in the γ5 region of γd KCs in naïve flies. Image acquisition is shuttered during color presentation. Right, quantification shows excitatory responses to blue, green lights and control (LED off). **d** and **e**. Pulsed delivery of colored light does not affect multisensory memory performance in the Congruent protocol (**d**) or Visual Retrieval (**e**) following multisensory learning. Asterisks denote significant differences (*P* < *0.05*). Data presented as mean ± standard error of mean (SEM). Individual data points displayed as dots correspond to independent experiments. **f** and **g**. *Top Left*, appetitive multisensory (color+odor, **f**) and unisensory (odor, **g**) training and odor imaging timelines. *Bottom left*, Imaging plane in γd KC somata. *Middle*, traces of CS+ and CS− odor-evoked activity. Inhibitory responses to CS+ and CS− odors (decrease in $-\Delta F/F_0$) were evident following either paradigm. *Right*, quantification. **h** and **i**. *Top Left*, appetitive multisensory (color+odor) training and odor imaging timelines. *Bottom left*, imaging plane in the γ5 (**h**) and γ1 (**i**) regions of γm KCs. *Middle*, traces of CS+ and CS− odor-evoked activity. Both γ1 and γ5 regions of γm KCs show excitation to CS+ and CS− odors (increases in $-\Delta F/F_0$). *Right*, quantification. For all traces and quantification, CS+ and CS− presentation involved 50:50

alternation of the odors as in all other experiments. Odor- or color-evoked activity traces show mean (solid line) with SEM (shadow). Horizontal dashed lines indicate baseline activity. Black line underneath traces marks 5 s odor exposure. In (**c**) gray vertical bar corresponds to the 0.75 s of color presentation (when image acquisition is shuttered) and box to the period (1.75 s) of quantification. Asterisks denote significant difference between averaged CS+ and CS− responses ($P < 0.05$). CS+ and CS− responses for each individual fly connected by dashed line. Groups compared using paired two-sided *t*-test (**a**,**b**, **f-i**), one-sample *t*-test (**c**), one-way ANOVA with Tukey's test (**d**), and unpaired two-sided *t*-test (**e**), exact *P* values and comparisons are given in Supplementary Information. N values for each experiment are: **a**, **b**, n = 20 flies; **c**, n = 20 flies for Green, n = 20 flies for Blue, n = 6 for LED off; **d**, **e**, n = 8; **f**, **g**, n = 22 flies; **h**, **i**, n = 16 flies.

Extended Data Fig. 4 Multisensory aversive learning enhances memory for the combined and individual odor and color cues.

a. *Top*, aversive training and testing timeline. *Bottom*, multisensory experimental conditions. Green and blue squares represent colors, light and dark gray represent OCT and MCH odors. Visual (V) learning; Olfactory (O) learning; Congruent (C) protocol; Incongruent (I) protocol; Olfactory Retrieval (OR); Visual Retrieval (VR). **b-d**. *Top*, training and testing timelines. *Bottom*, aversive memory with C protocol was significantly increased to that with I protocol both immediately (**b**) and 3 h (**c**) after training. Flies in the C protocol outperformed those tested 3 h after Olfactory (O) Learning. Only the C protocol generated significant 24 h memory performance (**d**). **e** and **f**. *Top*, training and testing timelines. *Bottom*, when tested immediately (**e**) multisensory memory retrieved with colors (VR) was markedly better than that following unisensory Visual learning (V). At 3 h (**f**) multisensory trained flies performed significantly better when their memory was retrieved with only Olfactory (OR) or Visual (VR) cues than flies trained with unisensory Olfactory (O) or Visual (V) learning. **g**. *Top left*, schematic of γd KCs. *Bottom left*, training and testing timeline with temperature shifting (dashed line). **g-i**. Blocking output of γd KCs during testing using MB607B-GAL4; UAS-*Shi*^{ts1} alters 3 h memory performance in the Congruent (**g**), Incongruent (**h**) and Olfactory Retrieval (**i**) protocols. **j**. *Top left*, schematic of γd KCs. *Bottom left*, training and

testing timeline with constant permissive temperature (dashed line). **j-l** Memory performance in MB607B-GAL4; UAS-*Shi*^{ts1} flies after Congruent (**j**), Incongruent (**k**), and Olfactory Retrieval (**l**) protocols was not affected when flies were trained and tested at 23 °C. Asterisks denote significant differences ($P < 0.05$). Data presented as mean ± standard error of mean (SEM). Individual data points displayed as dots correspond to independent experiments. **m** and **n**. *Top left*, unisensory aversive training and odor imaging protocol. *Bottom left*, imaging plane in the γ1 region (**m**) and γ5 (**n**) of γd KCs. *Middle*, traces of CS+ and CS− odor-evoked activity. uniisensory aversive training does not alter odor responses in either γd KC region. *Right*, quantification. **o** and **p**. *Top left*, unpaired aversive multisensory training and odor imaging protocol. Color+odor presentation was not contingent with shock. *Bottom left*, imaging plane in the γ1 region (**o**) and γ5 (**p**) of γd KCs. *Middle*, traces of CS+ and CS− odor-evoked activity. Unpaired multisensory aversive training does not alter odor responses in either γd KC region. *Right*, quantification. For all traces and quantification, CS+ and CS− presentation involved 50:50 alternation of the odors as in all other experiments. Odor-evoked activity traces show mean (solid line) with SEM (shadow). Horizontal dashed lines indicate baseline activity. Black line underneath traces marks 5 s odor exposure. Asterisks denote significant difference between averaged CS+ and CS− responses ($P < 0.05$). CS+ and CS− responses for each individual fly connected by dashed line. Groups compared using one-way ANOVA with Tukey's test (**b-d**, **g-l**), unpaired two-sided *t*-test (**e**, **f**), and paired two-sided *t*-test (**m-p**), exact *P* values and comparisons are given in Supplementary Information. N values for each experiment are: **b**, **g**, **i-l**, n = 8; **c**, **e**, **f**, **h**, n = 10; **d**, n = 12; **m**, **n**, n = 22 flies; **o**, **p**, n = 24 flies.

Extended Data Fig. 5 DPM and APL connectivity allows for multisensory stimulus binding.

a. A 2-dimensional dendrogram projection of DPM neuron neurites (shades of teal). The dorsal branch of the horizontal lobe's γ2-γ5 compartments are dark teal, the rest of the γ lobe projections are mid teal and those in the other lobes are lighter teal. Note: all but γ lobe neurites are downsized for visibility. Neurites with Strahler order <1 are pruned and connectivity is shown accordingly. Projections in the γ1 compartment are marked and split

in the γ2-5 compartments into dorsal and ventral branches. Marking is according to DAN connectivity (shaded areas). Synapses (spheres) are only marked in the γ lobe compartments. Connectivity structure shows inputs from γd KCs (yellow spheres) and outputs to γm KCs (gray spheres) colocalize on compartment specific branches of the DPM neurons. APL inputs (magenta) are distributed across both dorsal and ventral branches and concentrated around the branching point at the base of the vertical lobes in the α′1 compartment (asterisk). Note: Some putative parts of the dorsal γ lobe branch could not be allocated to a compartment due to lack of DAN input. One branch that bears γ KC synapses (hashtag) was identified as the β′ lobe branch that enters the MB close to the midline. It is likely that other DPM neurite tips that have γ KC synapses are artifacts where healing has merged free-floating γ lobe branchlets to DPM neurites innervating the other lobes. **b**. A 2-dimensional dendrogram projection of APL neuron neurites (magenta). Note: Neurites with Strahler order <1 were pruned and connectivity is shown accordingly. Compare with (**a**). PPL1-γ1pedc and PAM-γ5 DAN input synapses (red and green spheres respectively) are marked in the γ1 and γ5 compartments (shaded areas). Synapses to DPM neurons (teal spheres) colocalize with DAN input.

Extended Data Fig. 6 GABAergic APL neurons contribute to multisensory learning.

a and **b**. *Top*, Anatomy of APL neuron. Training and testing timelines with constant restrictive temperature (dashed line). *Bottom*, Adult-restricted RNAi knockdown of Dop2R in APL neurons using *tub*PGAL80^{ts};VT43924-GAL4.2 impaired 6 h Olfactory Retrieval memory performance after multisensory appetitive training (**a**), and reduced memory following unisensory Olfactory learning (**b**). **c** and **d**. Permissive temperature (23 °C) control experiments for (**a**) and (**b**). *Top*, training and testing timelines with temperature shift (dashed line). Memory performance was unaffected in *tub*PGAL80^{ts};VT43924-GAL4.2, UAS-Dop2R RNAi flies following Olfactory Retrieval (**c**) and Olfactory learning (**d**) when flies were raised, trained and tested at 23 °C. **e** and **f**. *Top*, training and testing timelines with temperature (dashed line). *Bottom*, 6 h Olfactory Retrieval of appetitive multisensory memory (**e**) and of memory following Olfactory learning (**f**) was not affected by adult-restricted RNAi knockdown of DopEcR in APL

neurons using *tub*PGAL80^{ts}; VT43924-GAL4.2. Asterisks denote significant differences ($P < 0.05$). Data presented as mean ± standard error of mean (SEM). Individual data points displayed as dots correspond to independent experiments. All groups compared using one-way ANOVA with Tukey's test, exact *P* values and comparisons are given in Supplementary Information. N values for each experiment are: **a**, **b**, n = 12; **c**, **d**, n = 8; **e**, **f**, n = 10.

Extended Data Fig. 7 DPM neurons mediate multisensory binding.

a. Permissive temperature (23 °C) control experiment for Fig. 5c,d. *Top*, training and testing timeline. *Bottom*, Multisensory memory performance evoked by Olfactory Retrieval of VT64246-GAL4; UAS-*Shi*^{ts1} flies was unaffected when flies were trained and tested at 23 °C. **b**. *Top*, training and testing timeline with temperature shifting (dashed line). *Bottom*, blocking output from DPM neurons during testing impaired the Visual Retrieval of multisensory memory. **c**. Permissive temperature (23 °C) control experiment for Extended Data Fig. 7b. *Top*, training and testing timeline. *Bottom*, Multisensory memory performance evoked by Visual Retrieval of VT64246-GAL4; UAS-*Shi*^{ts1} flies was unaffected when flies were trained and tested at 23 °C. **d**. *Top*, training and testing timeline with temperature shifting (dashed line). *Bottom*, blocking output from DPM neurons during unisensory training and testing did not affect Olfactory learning performance. **e**. *Top left*, schematic of γ KCs. **e-g**. *Top*, training and testing timeline with temperature shifting (dashed line). *Bottom*, blocking output of γ KCs during training using MB009B-GAL4; UAS-*Shi*^{ts1} did not affect 6 h Olfactory learning performance (**e**), while Olfactory Retrieval (**f**) and Visual Retrieval (**g**) memory performance was significantly impaired. **h** and **i**. Permissive temperature (23 °C) control experiments for Extended Data Fig. 7f,g. *Top*, training and testing timelines. *Bottom*, Multisensory memory performance evoked by Olfactory Retrieval (**h**) or Visual Retrieval (**i**) of MB009B-GAL4; UAS-*Shi*^{ts1} flies was unaffected when flies were trained and tested at 23 °C. **j** and **k**. *Left*, appetitive unisensory (odor) training and odor imaging timeline. Imaging planes in the γ5 (**j**) and γ1 (**k**) regions of the DPM neuron. *Right*, quantification. 6 h after olfactory training the γ5 (**j**) and γ1 (**k**) regions of DPM neurons showed indistinguishable excitatory responses to both CS+

and CS− odors. **k** and **l**. *Top*, appetitive multisensory (color+odor) or unisensory (odor) training and odor imaging timelines. Imaging plane in the γ5 region of DPM neurons. *Bottom*, quantification of γ5 responses of DPM neurons 1 h after both multisensory (**l**) and unisensory (**m**) training showed an enhanced excitatory response to the CS+ over the CS− odors. **n** and **o**. *Top*, appetitive multisensory (color+odor) or unisensory (odor) training and odor imaging timelines. Imaging plane in the γ1 region of the DPM neuron. *Bottom*, quantification of γ1 responses of DPM neurons 1 h after both multisensory (**n**) and unisensory (**o**) training showed no difference between the excitatory responses to the CS+ and the CS− odors. **p** and **q**. *Top*, imaging plane in the γ1 (**p**) and γ5 (**q**) region of γd KCs. *Bottom left* panels, a baseline recording in saline (300 s; not shown) was followed by perfusion of 100 μM 5-HT for 300 s then washout in saline for 300 s. Average traces for bath application and washout are shown. *Right*, quantification *shows* excitatory responses to 5-HT application in both regions (increased mean $-\Delta F/F_0$), in comparison to the baseline (pre) and after washout. **r**. *Top left*, schematic of γd KCs. *Bottom left*, training and testing timeline. *Right*, 6h Olfactory learning performance is unaffected in flies with RNAi knockdown of 5-HT2A with MB607B-GAL4. Asterisks denote significant differences ($P < 0.05$). Data presented as mean ± standard error of mean (SEM). Individual data points displayed as dots correspond to independent experiments. **s**. *Left*, appetitive multisensory (color+odor) training and color imaging timeline. Imaging plane in the γ5 region of γd KC axons. *Middle*, traces of CS+ and CS− color-evoked activity. The γ5 region of γd KC axons showed excitatory responses to both CS+ and CS− colors. *Right*, quantification of color-evoked responses. For trace in (**s**) and all quantifications, CS+ and CS− presentation involved 50:50 alternation of the odors and/or colors as in all other experiments. Color-evoked activity traces show mean (solid line) with SEM (shadow). In (**s**) vertical gray bar corresponds to the 0.75 s color presentation (when image acquisition is shuttered) and box to the 1.75 s period of quantification. Horizontal dashed lines indicate baseline activity. Asterisks denote significant difference between averaged CS+ and CS− responses ($P < 0.05$). CS+ and CS− responses for each individual fly connected by dashed line. Groups compared using one-way ANOVA with Tukey's test (**a-e**, **g-i**, **r**), Kruskal–Wallis *H*-test with Dunn's test (**f**), paired two-sided *t*-test (**j-o**, **s**), and repeated measures one-way ANOVA with Tukey's test (**p,q**), exact *P* values

and comparisons are given in Supplementary Information. N values for each experiment are: **a-c**, **r**, n = 8; **d**, n = 12; **e-i**, n = 8; **j**, **k, n**, n = 18 flies; **l**, **m**, **o**, **s**, n = 16 flies; **o**, **p**, n = 10.

Supplementary information

Reporting Summary

Supplementary Table 1

n values, descriptive statistics and statistical analyses with exact *P*-values for Figs. 1–5.

Supplementary Table 2

n values, descriptive statistics and statistical analyses with exact *P*-values for Extended Data Figs. 1–7.

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Okray, Z., Jacob, P.F., Stern, C. *et al.* Multisensory learning binds neurons into a cross-modal memory engram. *Nature* **617**, 777–784 (2023). https://doi.org/10.1038/s41586-023-06013-8

- Received: 07 July 2022
- Accepted: 24 March 2023
- Published: 26 April 2023
- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-06013-8

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

How Rosalind Franklin’s story was rewritten

- Benjamin Thompson
- Shamini Bundell

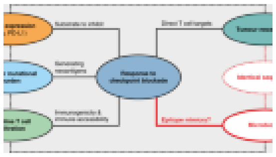
Nature Nature Podcast 26 Apr 2023

How flies remember a rich experience from its individual components

Nature Research Briefing 26 Apr 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06013-8

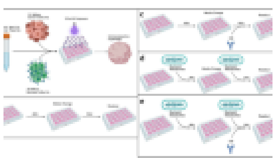
| Section menu | Main menu |



Tumour neoantigen mimicry by microbial species in cancer immunotherapy

06 April 2021

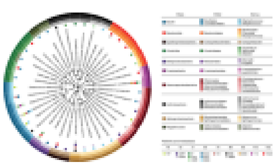
Maximilian Boesch, Florent Baty, … Martin H. Brutsche



Immuno-reactive cancer organoid model to assess effects of the microbiome on cancer immunotherapy

15 June 2022

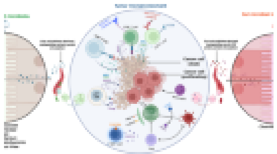
Ethan Shelkey, David Oommen, … Shay Soker



Identification of bacteria-derived HLA-bound peptides in melanoma

17 March 2021

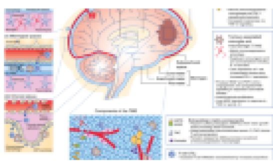
Shelly Kalaora, Adi Nagler, … Yardena Samuels



Potential links between the microbiota and T cell immunity determine the tumor cell fate

24 February 2023

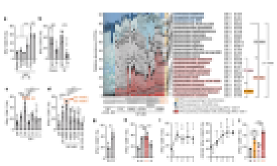
Amir Hossein Mohseni, Sedigheh Taghinezhad-S, … Dan Li



Harnessing the immune system in glioblastoma

05 November 2018

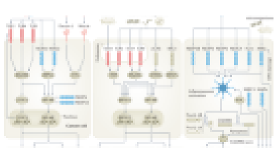
Nicholas F. Brown, Thomas J. Carter, … Paul Mulholland



A defined commensal consortium elicits CD8 T cells and anti-cancer immunity

23 January 2019

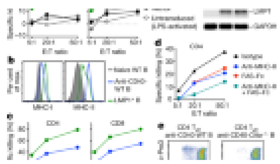
Takeshi Tanoue, Satoru Morita, … Kenya Honda



Context-dependent functions of pattern recognition receptors in cancer

30 March 2022

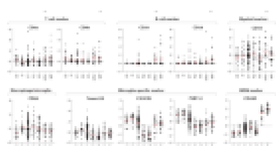
Si Ming Man & Brendan J. Jenkins



Mechanism of EBV inducing anti-tumour immunity and its therapeutic use

23 December 2020

Il-Kyu Choi, Zhe Wang, … Baochun Zhang



The immune landscape of common CNS malignancies: implications for immunotherapy

11 June 2021

Martina Ott, Robert M. Prins & Amy B. Heimberger

- Article
- Open Access
- Published: 17 May 2023

Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma

- Reza Naghavian,
- Wolfgang Faigle,
- Pietro Oldrati,
- Jian Wang,
- Nora C. Toussaint,
- Yuhan Qiu,
- Gioele Medici,
- Marcel Wacker,
- Lena K. Freudenmann,
- Pierre-Emmanuel Bonté,
- Michael Weller,
- Luca Regli,
- Sebastian Amigorena,
- Hans-Georg Rammensee,
- Juliane S. Walz,
- Silvio D. Brugger,
- Malte Mohme,
- Yingdong Zhao,
- Mireia Sospedra,
- Marian C. Neidert &
- …
- Roland Martin

Show authors

Nature **volume 617**, pages 807–817 (2023)

- 16k Accesses

- 67 Altmetric

- Metrics details

Abstract

Microbial organisms have key roles in numerous physiological processes in the human body and have recently been shown to modify the response to immune checkpoint inhibitors[1,2]. Here we aim to address the role of microbial organisms and their potential role in immune reactivity against glioblastoma. We demonstrate that HLA molecules of both glioblastoma tissues and tumour cell lines present bacteria-specific peptides. This finding prompted us to examine whether tumour-infiltrating lymphocytes (TILs) recognize tumour-derived bacterial peptides. Bacterial peptides eluted from HLA class II molecules are recognized by TILs, albeit very weakly. Using an unbiased antigen discovery approach to probe the specificity of a TIL $CD4^+$ T cell clone, we show that it recognizes a broad spectrum of peptides from pathogenic bacteria, commensal gut microbiota and also glioblastoma-related tumour antigens. These peptides were also strongly stimulatory for bulk TILs and peripheral blood memory cells, which then respond to tumour-derived target peptides. Our data hint at how bacterial pathogens and bacterial gut microbiota can be involved in specific immune recognition of tumour antigens. The unbiased identification of microbial target antigens for TILs holds promise for future personalized tumour vaccination approaches.

Main

The treatment of cancer includes the three classical ‘pillars’: surgery, radiotherapy and chemotherapy. With cancer immunotherapy, a fourth pillar emerged[3]. The introduction of immune checkpoint inhibition led to advances in survival rates in several cancers and demonstrated that unleashing immune effector mechanisms can result in the elimination of tumour cells[4]. In parallel, numerous approaches are being explored to go a step further and vaccinate patients with tumour neoantigens to induce tumour-specific T cells that mount an effective immune response against the tumour[5]. Several lines of evidence indicate that tumour vaccination will be feasible, particularly in ‘hot’ tumours with high numbers of mutations and strong immune cell infiltrates[6,7]. However, numerous hurdles still need to be overcome

including the relatively low immunogenicity of many tumour antigens, which, if not mutated, do not elicit strong antitumour responses, as T cells with reactivity against self-antigens are eliminated by thymic central tolerance[8]. Hence, the more foreign an antigen 'looks' to T cells, the more likely it will induce a strong immune response. In agreement with this, reports about patients, whose tumour either shrank substantially during an infection with a bacterial[9] or viral pathogen[10] or disappeared, indicated that protective immune responses against pathogens may also target the tumour, most likely by T cell cross-reactivity against tumour-derived and pathogen-derived antigens[11]. Supporting this notion, several studies have demonstrated improved responses to immune checkpoint inhibitors in the presence of certain gut bacteria[1,2]. Adoptive transfer of *Bacteroides fragilis*-specific T cells to germ-free mice restored the response to anti-CTLA4 (ref. [2]). Indeed, antitumour responses can be elicited by cross-reactivity between commensal bacteria and tumour antigens[12]. Furthermore, $CD8^+$ T cells can cross-recognize a peptide from *Enterococcus hirae* bacteriophage and a tumour antigen by molecular mimicry[13]. One approved therapy for bladder cancer, an extract of the *Mycobacterium bovis* strain Bacille Calmette Guerin, which is locally instilled into the bladder, has been shown to induce T helper 1 (T_H1) $CD4^+$ T cell responses against the tumour and to provide long-term protection in mice[14,15]. The importance of $CD4^+$ T cells in driving antitumour responses are highlighted by showing that melanoma-derived neoepitopes are recognized by $CD4^+$ T cells[16], and neoantigen vaccinations in melanoma and glioblastoma primarily activate $CD4^+$ T cells[17,18,19].

As another line of evidence hinting at a possible involvement of pathogens or commensal bacteria in antitumour responses, several tumours harbour bacteria[20]. Melanoma metastases even present bacterial peptides on their HLA class I and class II molecules[21], indicating that these are available for T cell recognition and thus could be used for tumour vaccination. Intratumoural bacteria may influence tumour outcome positively and negatively in several ways, including direct effects on transcriptional pathways of inflammation, tumour growth and immune cell infiltration[22,23]. Regarding antitumour immunity, microbiota participate in shaping the adaptive immune repertoire by stimulation via pathogen-associated molecular patterns, by immune-modulating metabolites and, most

interestingly in the context of tumour vaccination, by inducing cross-reactive T cell responses[11].

Here we pursued several experimental paths to address whether and how microbial antigens may contribute to immune reactivity against glioblastoma. We first examined the glioblastoma immunopeptidomes of 19 patients for HLA class II-bound bacterial peptides, and next whether TILs and TIL-derived $CD4^+$ T cell clones (TCCs) in one of the patients react against HLA-bound bacterial peptides and tumour antigens. We then used an unbiased antigen discovery approach to identify targets derived from viruses, bacteria and bacterial gut microbiota, which we refer to as gut microbiota from now on. TCC88, which is directed to a glioblastoma neoantigen, not only recognizes multiple other glioblastoma-derived peptides but also responds strongly to a broad spectrum of bacterial-derived and gut microbiota-derived targets, which also elicit cross-reactive T cell responses against tumour targets in bulk TILs and even peripheral blood memory T cells.

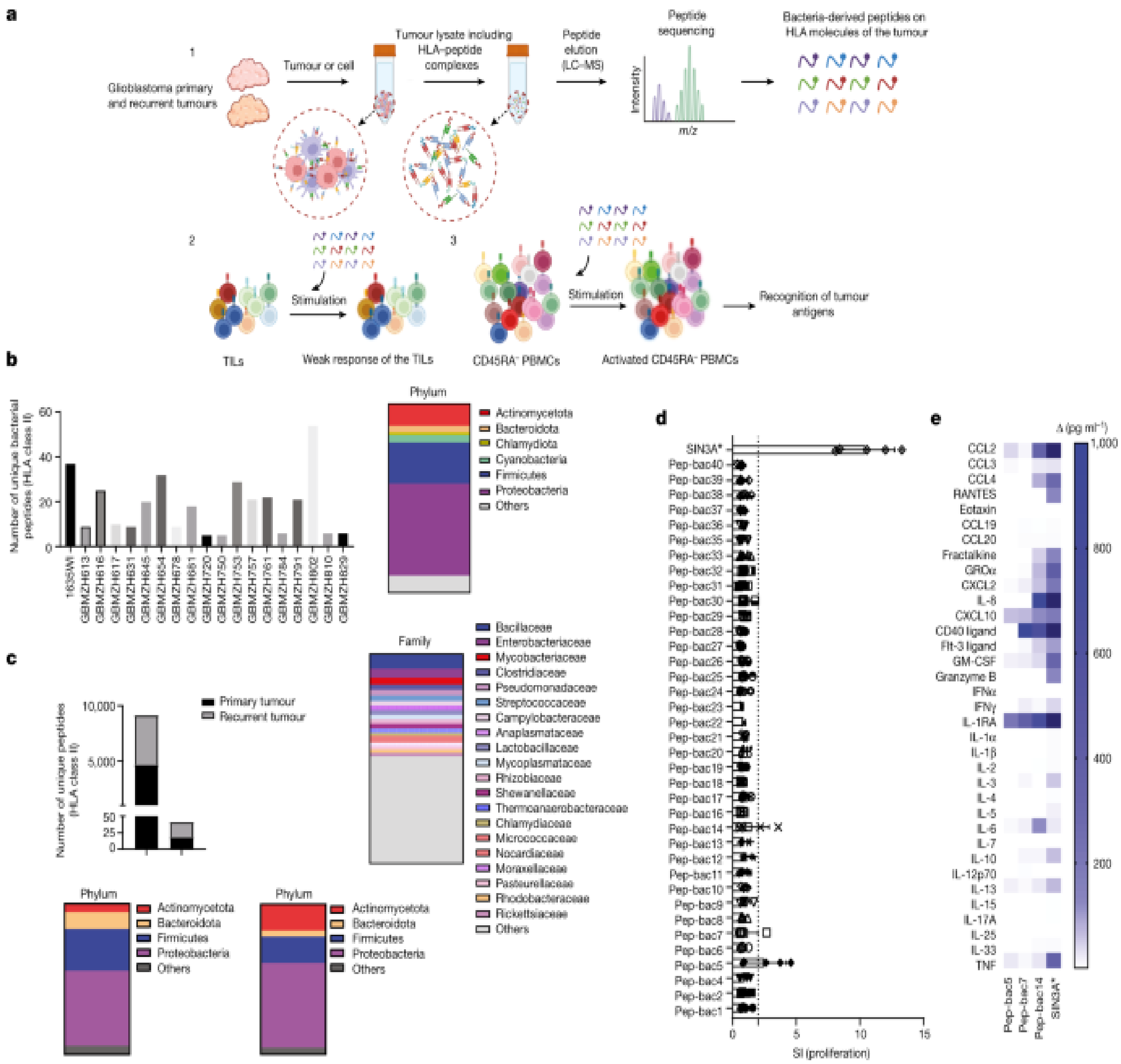
Microbial peptides in HLA class II of glioblastoma

A study has reported that tumours contain distinct microbial communities including glioblastoma[20]. Furthermore and supporting that bacterial antigens within tumours could in principle elicit intratumoural immune responses, another study has found bacterial peptides in the HLA class II immunopeptidome of melanoma cells[21]. Although it appears easier that bacterial peptides can be presented by tumour cells or local antigen-presenting cells (APCs) in the skin, we examined this topic here in glioblastoma. We first analysed the HLA class I-associated and HLA class II-associated immunopeptidomes of 19 primary or recurrent glioblastomas (Supplementary Tables 1 and 2) and annotated these for human proteins, tumour-specific antigens (that is, mutated tumour antigens) and also for sequences from bacteria (UniProtKB/Swiss-Prot protein database) and viruses with tropism for humans (Swiss-Prot-20210305) (Fig. 1a). We identified not only multiple tumour-associated antigens, that is, self-proteins overexpressed in the tumour, and tumour-specific antigens on both HLA class I and class II molecules (data not shown) but also between 5 and 54 unique HLA class II-bound bacterial sequences per patient and tumour, but

no viral peptides (Fig. 1b, left, and Supplementary Table 3). As controls, we analysed the immunopeptidomes from brain lesions of patients with multiple sclerosis (MS) and healthy brain tissue[24] (dataset identifier PXD019643). We found between 10 and 45 bacterial hits from brain tissues of three patients with MS and six healthy donors (Extended Data Fig. 1a and Supplementary Table 4). Seven peptides from the control cohorts exactly matched peptides in glioblastoma samples, and we therefore subtracted these from the bacterial peptides in glioblastoma. Tumour-derived bacterial peptides originated from 255 bacterial proteins and 199 unique bacteria mainly from the phyla Proteobacteria and Firmicutes, with Enterobacteriaceae, Bacillaceae and Mycobacteriaceae among the top families from each phylum (Fig. 1b, right, and Supplementary Table 3). The composition and number of bacterial peptides differed between patients, although several shared HLA alleles (Supplementary Table 2). Furthermore, we generated tumour cell lines from six of the patients and analysed the immunopeptidomes (Supplementary Table 1). The tumour cell lines presented similar numbers of HLA class II-bound bacterial peptides, which we refer to here as immunopeptidome-derived bacterial peptides (IPdBPs). These observations indicate that bacterial peptides exist in the tumour cells, are processed and loaded onto HLA molecules and form part of the IPdBPs. However, the number of IPdBPs of both tumour tissue and cell lines was low (Supplementary Table 3). In addition, both tumour and isolated cell lines showed normal expression of HLA class II in patient 1635WI (Extended Data Fig. 2). We can therefore not rule out that at least part of the IPdBPs are presented by APCs in the tumour. We further conducted 16S rRNA gene sequencing on fresh frozen tumour tissues of ten patients with glioblastoma. Similar to immunopeptidome analyses, bacterial loads in the tumours were low and Proteobacteria accounted for the majority of them (Extended Data Fig. 1b and Supplementary Tables 5 and 6). We then compared the 16S rRNA results with bacterial sources of the IPdBPs in each patient. We found at least one exact match for bacterial species in five out of ten patients (Supplementary Table 7). As the resolution at the species level is limited due to high similarities of 16S rRNA sequences between certain bacteria[21], we examined the data at the next taxonomic level, bacterial genus, and found matches in seven patients. These data suggest the presence of both bacterial DNA as well as peptides in glioblastoma. Next, we assessed the possible immune recognition of IPdBPs in detail in patient

1635WI, whom we had previously treated with a personalized neoantigen vaccination and checkpoint inhibitors in addition to standard therapy[25]. We identified 18 and 23 unique IPdBPs from primary and recurrent tumour, respectively (Fig. 1c, top, and Supplementary Tables 8 and 9), and four in both tumours. Peptides of the primary and recurrent tumours were derived from 15 and 21 bacterial proteins (Supplementary Table 8). Consistent with the immunopeptidome data of the entire group of patients, Proteobacteria and Firmicutes were most common (Fig. 1c, bottom). Furthermore, we compared the deduced core binding motifs and HLA-DR anchor positions between bacterial and self-antigen or tumour-antigen-derived peptides within this patient (Extended Data Fig. 1c). Amino acid preferences at HLA-DR anchor positions were similar between bacterial and self/tumour peptides, but different at non-anchor positions (Extended Data Fig. 1c).

Fig. 1: Elution of HLA class II-presented bacterial peptides from 19 glioblastomas and the stimulatory effect of IPdBPs on bulk TILs.



a, Schematic of the immunopeptidome analyses of glioblastoma tissue, from which eluted peptides were annotated using the UniProt bacteria database (1). TILs or peripheral memory T cells were stimulated with IPdBPs (2 and 3). LC–MS, liquid chromatography–mass spectrometry; PBMC, peripheral blood mononuclear cell. The images in panel **a** were created using BioRender (https://biorender.com). **b**, Number of unique IPdBPs in patients with glioblastoma (*n* = 19) (data from patient 1635WI contains both primary and recurrent tumours together), with 344 unique IPdBPs identified (left). These peptides originated from 255 and 199 unique proteins and bacteria, respectively (3 technical replicates). The taxonomy of these bacteria shows that they primarily stem from Proteobacteria and Firmicutes at the phylum level (right). **c**, IPdBPs from primary and recurrent tumours (patient

1635WI) (top). As a control, peptidome data were also annotated using human protein database to identify eluted self-peptides. As shown, most of HLA class II-bound peptides are self-(human) peptides. IPdBPs of primary (bottom left, $n = 18$) and recurrent (bottom right, $n = 23$) tumours were separately annotated to their source bacteria at the phylum level. **d**, Bulk TILs (patient 1635WI) were stimulated with IPdBPs (See Supplementary Table 8), each in five replicate wells, and proliferation was measured after 5 days via ^{3}H-thymidine incorporation. Proliferation is shown as a stimulatory index (SI), and peptides with SI ≥ 2 are considered stimulatory (the dotted line shows SI = 2 throughout the paper). The data are expressed as the mean value of all wells ± s.e.m. SIN3A* peptide was used as a positive control as TILs recognize this tumour antigen. T cells with APCs without any stimuli were used as a negative control. **e**, Heatmap of cytokine and chemokine secretion of TILs stimulated with three IPdBPs, which weakly stimulated proliferation. The secretion of each cytokine and chemokine was calculated by subtracting the corresponding cytokine and chemokine amount (pg ml^{-1}) of the negative control. TILs with APCs without any stimuli were used as a negative control. Also see Extended Data Fig. 1.

Source data

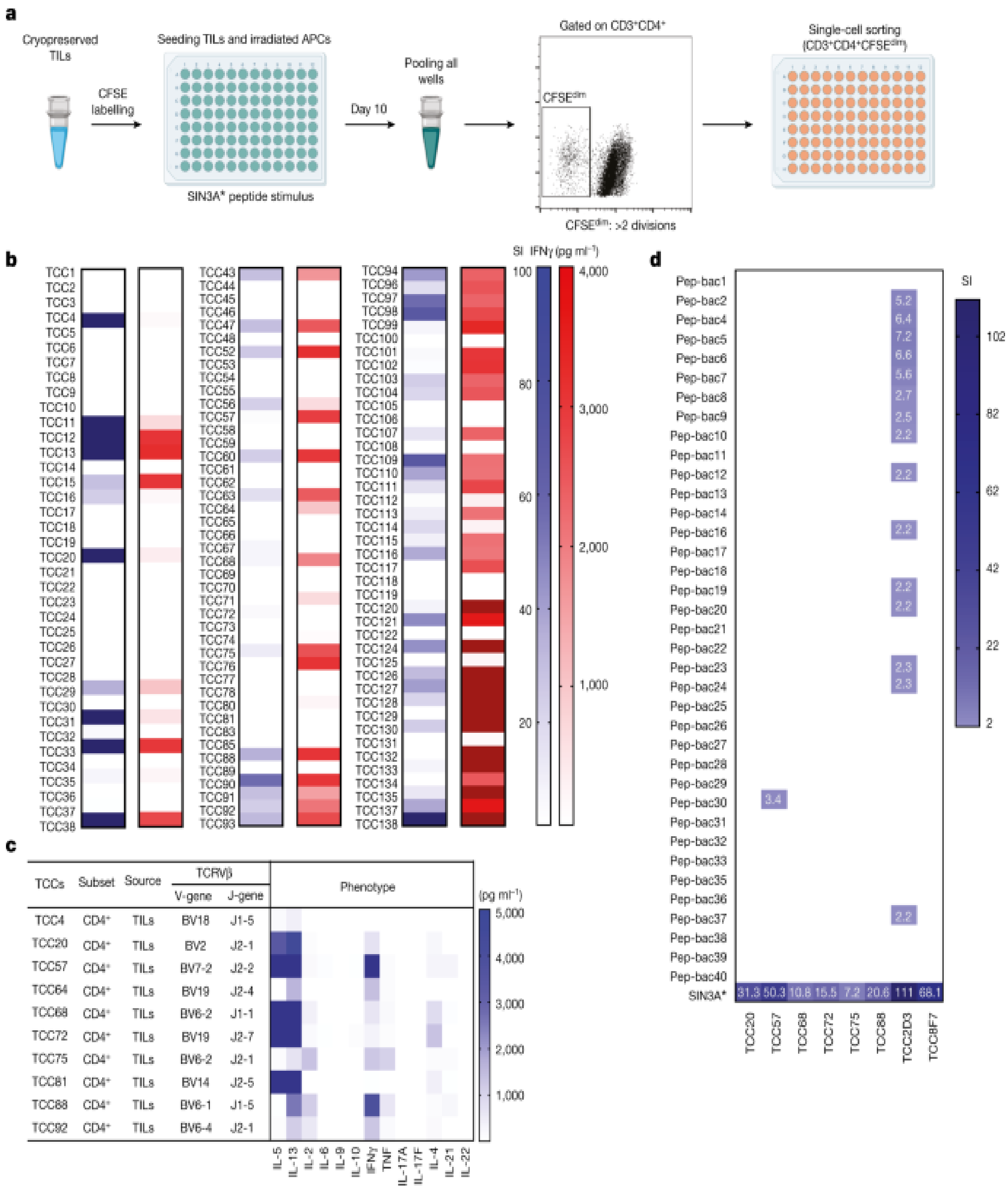
We further wanted to know whether IPdBPs are involved in the intratumoural immune response and therefore tested bulk TILs from the recurrent glioblastoma (patient 1635WI) with all 37 unique IPdBPs. Compared with a tumour neoantigen derived from the paired amphipathic helix protein SIN3A, peptide SINA3A* 300–312 (QPVEFNHAIHYVN; '*' marks all mutated proteins throughout the paper), which was part of the tumour vaccine, the proliferative response of TILs to the 37 IPdBPs was very weak (Fig. 1d and Supplementary Table 10); however, they secreted pro-inflammatory cytokines including IFNγ, TNF and GM-CSF upon stimulation with several IPdBPs (Fig. 1e).

Tumour-reactive CD4^{+} TCCs

The above data indicate that bacteria-derived peptides may be involved in intratumoural immune responses. As bulk TIL responses to IPdBPs were weak, we isolated TCCs from the TILs that recognize a tumour antigen with

the goal of testing them for possible recognition of IPdBPs. On the basis of the response of bulk TILs to a strongly stimulatory peptide from the personalized vaccine cocktail[25], we used a SIN3A* peptide to stimulate bulk TILs (Fig. 2a and Extended Data Fig. 3a). TCCs were then generated from proliferating $CD4^+$ (carboxyfluorescein succinimidyl ester $(CFSE)^{dim}$) T cells (Fig. 2a). Of $CD4^+$ TCCs, 70 out of 124 responded to the SIN3A* peptide by proliferation or production of IFNγ (Fig. 2b). TCCs were characterized further for TCR expression and functional phenotypes. Most showed a polyfunctional T_H1/T_H2 phenotype with production of IFNγ, IL-5 and IL-13 after stimulation via anti-CD2, anti-CD3 and anti-CD28 antibodies (Fig. 2c). Next, we tested several of these tumour-reactive TCCs with IPdBPs. TCC57 and TCC2D3 responded to one and several IPdBPs; however, again relatively weakly compared with the SIN3A* peptide (Fig. 2d).

Fig. 2: Isolation and characterization of tumour-specific TCCs by stimulating TILs with SIN3A* peptide.



a, Bulk TILs were labelled with CFSE and stimulated with irradiated autologous PBMCs as APCs pulsed with SIN3A* peptide. After 10 days, 60 replicate wells were pooled and $CFSE^{dim}CD3^{+}CD4^{+}$ T cells were single-cell sorted and expanded using PHA and human IL-2. The images in panel **a** were created using BioRender (https://biorender.com). **b**, All isolated TCCs were tested against the cognate antigen (SIN3A* peptide) for proliferation

and IFNγ production. After 3 days, proliferation was measured using ^{3}H-thymidine incorporation (SIs), and ELISA was used to measure IFNγ secretion. The response of each TCC was compared with the stimulation of the same TCC with APCs without peptide. **c**, TCRB V-gene and J-gene and functional phenotypes of several of the TCCs. **d**, Several TIL-derived TCCs with reactivity to tumour antigen (SIN3A*) were tested for reactivity against IPdBPs. TCCs were stimulated with peptides and irradiated PBMCs as APCs for 3 days. Proliferation was measured using ^{3}H-thymidine incorporation. SIN3A* peptide served as the positive control. Mean SIs are shown in a heatmap format. SI < 2 (non-stimulatory) are shown with an empty field, and only SI ≥ 2 for stimulatory peptides are shown. Also see Extended Data Fig. 3.

Source data

TCCs cross-recognize bacterial peptides

We next investigated the recognition of foreign antigens by intratumoural lymphocytes from a different angle, that is, not starting from the immunopeptidome, but with an unbiased antigen discovery methodology (Extended Data Fig. 3b), which we had previously used to identify the targets for autoreactive TCCs from the peripheral blood, cerebrospinal fluid and brain tissue of patients with MS[26,27], and in viral and bacterial infections[28,29]. It involves testing TCCs with synthetic combinatorial peptide libraries in the positional scanning format (ps-SCL). Individual ps-SCL mixtures are characterized by one defined amino acid in a determined position of a decamer, for example, l-alanine in position 1, whereas the remaining nine positions are randomized (Extended Data Fig. 3b). As a preparatory step[30], we characterized the HLA-DR restriction of several TCCs using irradiated Epstein–Barr virus (EBV)-transformed bare lymphocyte syndrome (BLS) B cells transfected with single autologous HLA-DR molecules including DRB1*03:01, DRB1*04:02, DRB3*02:02 or DRB4*01:01. Most CD4^{+} TCCs robustly recognized the SIN3A* peptide when presented by HLA-DRB3*02:02 (Extended Data Fig. 4a). As previously shown for virus-specific and autoantigen-specific TCCs[31,32], TCC88 showed cross-restriction by responding to the cognate peptide

presented on all four HLA-DR molecules but responded strongest to SIN3A*–HLA-DRB3*02:02 (half-maximal effective concentration (EC_{50}) = 0.013 μM) (Extended Data Fig. 4a). After testing a TCC with the 200 ps-SCLs, the experimental data (proliferation or cytokine release) were transformed into a scoring matrix that depicts the stimulatory score for each l-amino acid in each position of a decamer peptide[29] (Extended Data Fig. 3b). On the basis of the assumption that each amino acid contributes independently and additively to T cell recognition, the scoring matrix and a dedicated search algorithm allow scanning of any database to identify target peptides for the respective TCC in an unbiased way[29]. Peptides with predicted high stimulatory scores, for example, with higher than 80% of the maximal score, are then selected, synthesized and tested. Previously, for virus-specific, bacteria-specific and autoreactive TCCs, we had demonstrated that there is a strong relationship between the predicted stimulatory score and T cell recognition[29,33]. Among the above SIN3A*-specific TCCs, TCC88 responded well to the ps-SCL when these were presented by BLS cells transduced with HLA-DRA1*01:01 and HLA-DRB3*02:02 using IFNγ production as T cell activation readout (Extended Data Fig. 4b). Most of the other TCCs did not respond to ps-SCL (data not shown).

We next created a scoring matrix[29] using the median IFNγ values of three independent experiments and transforming these to $\log_{10}$ values (Extended Data Fig. 4c). Following the in silico predictions, over 300 peptides in the range between 60% and 100% of the maximum theoretical score (%) were synthesized and tested to assess and reconfirm the accuracy of the predictions. Of these, TCC88 recognized more than 80% of peptides within the score range of 90–100% of the maximum score (Extended Data Fig. 4d).

For the identification of putative targets of TCC88, we searched several large databases including all human proteins, the translated transcriptome of the primary and recurrent autologous glioblastoma, all viruses, which are annotated for infecting humans, as well as UniProt-deposited bacteria and human gut microbiota[34] (see Methods) (Fig. 3a). The numbers of unique 10-mers within each of these databases are shown in Fig. 3a. Combined, these represent more than 5×10^9 decamer peptides. We synthesized 542 peptides (see Methods) with high predicted stimulatory scores (Fig. 3a) and tested

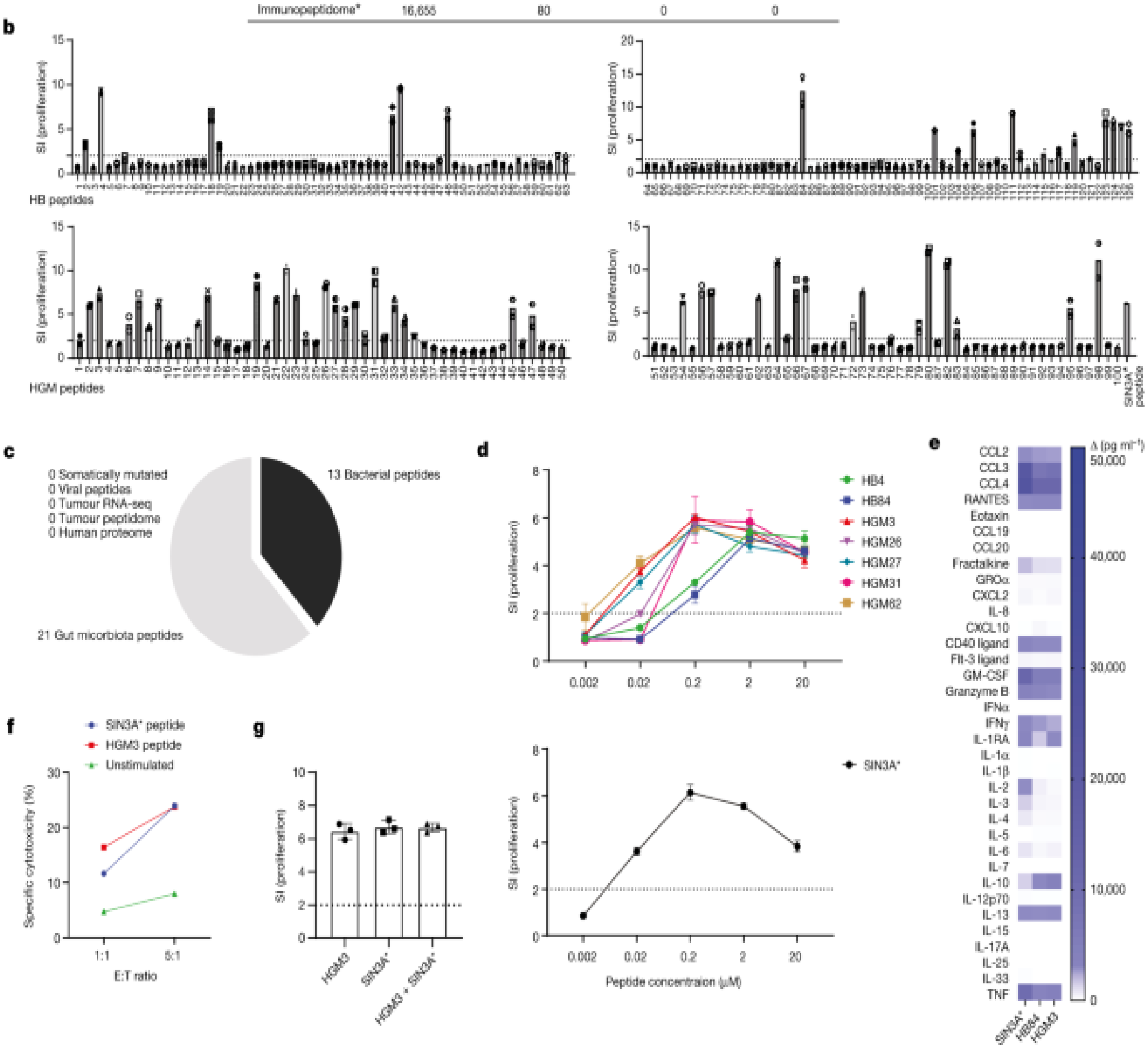
them with TCC88. As previously shown for other TCCs[26,27,28,29,32], TCC88 recognized multiple peptides from different sources including the autologous tumour transcriptome, the human proteome, viruses, bacteria and gut microbiota (Fig. 3a and Supplementary Table 10). Interestingly, 57.7% of all stimulatory peptides were derived from bacteria and gut microbiota, and 29.3% from tumour antigens containing tumour-associated antigens derived from the autologous glioblastoma, whereas peptides derived from other human proteins and viruses represented only minor fractions (Fig. 3a). Following our above observation, we were particularly interested in the bacteria-derived and microbiota-derived peptides that stimulated TCC88. We tested 126 bacteria-derived and 100 gut microbiota-derived peptides. The TIL-derived, SIN3A*-specific TCC88 recognized 21 bacterial and 42 gut microbiota-derived peptides (Fig. 3a,b) and 34 of these equally well or even better at a single-antigen concentration than the mutated tumour peptide, whereas peptides from other sources did not show this effect (Fig. 3b,c). The recognition with high functional avidity by TCC88 was confirmed by dose titration of several bacteria-derived and microbiota-derived peptides in the context of HLA-DRB3*02:02 (EC_{50} ~ 0.017, 0.018 and 0.019 μM for the peptides HGM62 (*Sedimentibacter*), HGM3 (*Victivallis*) and HGM27 (*Prevotella*), respectively), which is similar to SIN3A* (EC_{50} ~ 0.018 μM) (Fig. 3d). Despite the remarkable cross-recognition of multiple tumour-derived, bacteria-derived and microbiota-derived peptides, TCC88 did not respond to a large number of peptides (260 peptides) covering the whole-protein sequence of several brain and myelin, pancreatic and skin autoantigens (Extended Data Fig. 4e and Supplementary Table 11). The response of TCC88 to the unmutated SIN3A peptide confirmed that it weakly recognizes the wild-type antigen (Extended Data Fig. 4f). Upon activation with cognate antigen or bacteria-derived and microbiota-derived peptides, TCC88 shows a pro-inflammatory phenotype and secretes IFNγ, GM-CSF, TNF and granzyme B (Fig. 3e). The release of CCL3 (also known as MIP1α) and CCL4 (also known as MIP1β) indicates that activated TCC88 may recruit other immune cells such as $CD8^+$ T cells[35] (Fig. 3e). Although the cytokine and chemokine responses of TCC88 to different antigens were overall comparable, differences were seen, that is, higher IL-2 and fractalkine (also known as CX3CL1) levels upon stimulation with SIN3A* and increased levels of IL-10 and CXCL10 (also known as IP10) with bacteria-derived and microbiota-derived peptides (Fig.

3e). Furthermore, cytotoxic activity against autologous glioblastoma cells was comparable after stimulation with SIN3A* or HGM3 peptides (Fig. 3f). Owing to slow in vitro growth of patient-derived tumour cells, only HGM3 peptide was tested as it has the same EC_{50} as the SIN3A* peptide for the clone. Hence, TCC88 strongly cross-recognizes tumour antigens as well as bacteria-derived and microbiota-derived peptides, which are mainly derived from the phyla Firmicutes, Proteobacteria and Bacteroidota (Extended Data Fig. 5a). In line with previous data and other clones[26,32], we were not able to discern common amino acid patterns in stimulatory peptides when compared with non-stimulatory peptides (Extended Data Fig. 5b). Among stimulatory peptides are peptides from several species including *B. fragilis* (HGM32 peptide), which had previously been shown to be important for immune checkpoint inhibitor therapies[2] (Supplementary Table 10). Finally, stimulating TCC88 with both HGM3 and SIN3A* peptides did not have synergistic effects (Fig. 3g).

Fig. 3: Bacteria-derived and microbiota-derived peptides can stimulate TCC88 as well as the tumour antigen.

a

Databases	Number of unique 10-mer peptides	Number of synthesized peptides (80–100% of maximum score)	Number of stimulatory peptides	Stimulatory peptides within each database (%)
Human proteome	11,715,130	18	4	16.6
Gut microbiota	3,449,587,756	100	42	42
Human viruses	11,715,215	86	10	11.9
Bacteria	1,605,719,804	126	21	16.7
Tumour RNA-seq	8,924,646	132	32	24.2
Immunopeptidome*	16,655	80	0	0



a, Peptides with the highest predicted stimulatory scores were identified by searching various databases (see Methods). Numbers of unique 10-mer peptides within each database, synthesized peptides, TCC88 stimulatory peptides and the corresponding percentage of stimulatory peptides within each database are shown. Proliferation was measured using ^{3}H-thymidine incorporation (see Supplementary Table 10). *The immunopeptidome is annotated using human proteins in which 11 peptides score over 80%. Hence, the rest of the synthesized peptides were chosen from 60–80% of maximum score. RNA-seq, RNA sequencing. **b**, TCC88 was stimulated with

bacteria-derived (top) and microbiota-derived (bottom) peptides for 3 days and proliferation was measured by ^{3}H-thymidine incorporation (data represent the mean SI value of two replicate wells). HB, UniProt bacteria database; HGM, human gut microbiota[34]. **c**, Pie chart displaying the number of peptides from all databases that stimulated TCC88 equally well or better than the cognate antigen (SIN3A* tumour antigen) at a concentration of 10 μM. **d**, TCC88 was tested with several strongly stimulatory bacteria-derived and microbiota-derived and SIN3A* peptides at decreasing concentrations including 20, 2, 0.2, 0.02 and 0.002 μM. Irradiated BLS-DRB3*02:02 was pulsed with the peptides and used to stimulate TCC88 (data represent the mean SI value of five replicate wells ± s.e.m.). **e**, Supernatants were collected from TCC88 after stimulation with SIN3A*, HGM3 and HB84 peptides. Cytokine and chemokine secretion of TCC88 was examined and compared upon stimulation with these three different antigens. A specific amount of each cytokine and chemokine (pg ml^{-1}) was calculated by subtracting the cytokines in the negative control cultures (no peptide control). **f**, Cytotoxicity of TCC88 against the autologous tumour cell line after stimulation with 2 μM of SIN3A* and HGM3 (data represent the mean percentage of two replicate wells). E:T, effector to target. **g**, TCC88 was tested with HGM3 (10 μM) or SIN3A* (10 μM) peptides or both together (each 10 μM), and ^{3}H-thymidine incorporation was measured after 3 days (data represent the mean SI value of three replicate wells ± s.e.m.). Also see Extended Data Figs. 3 and 4.

Source data

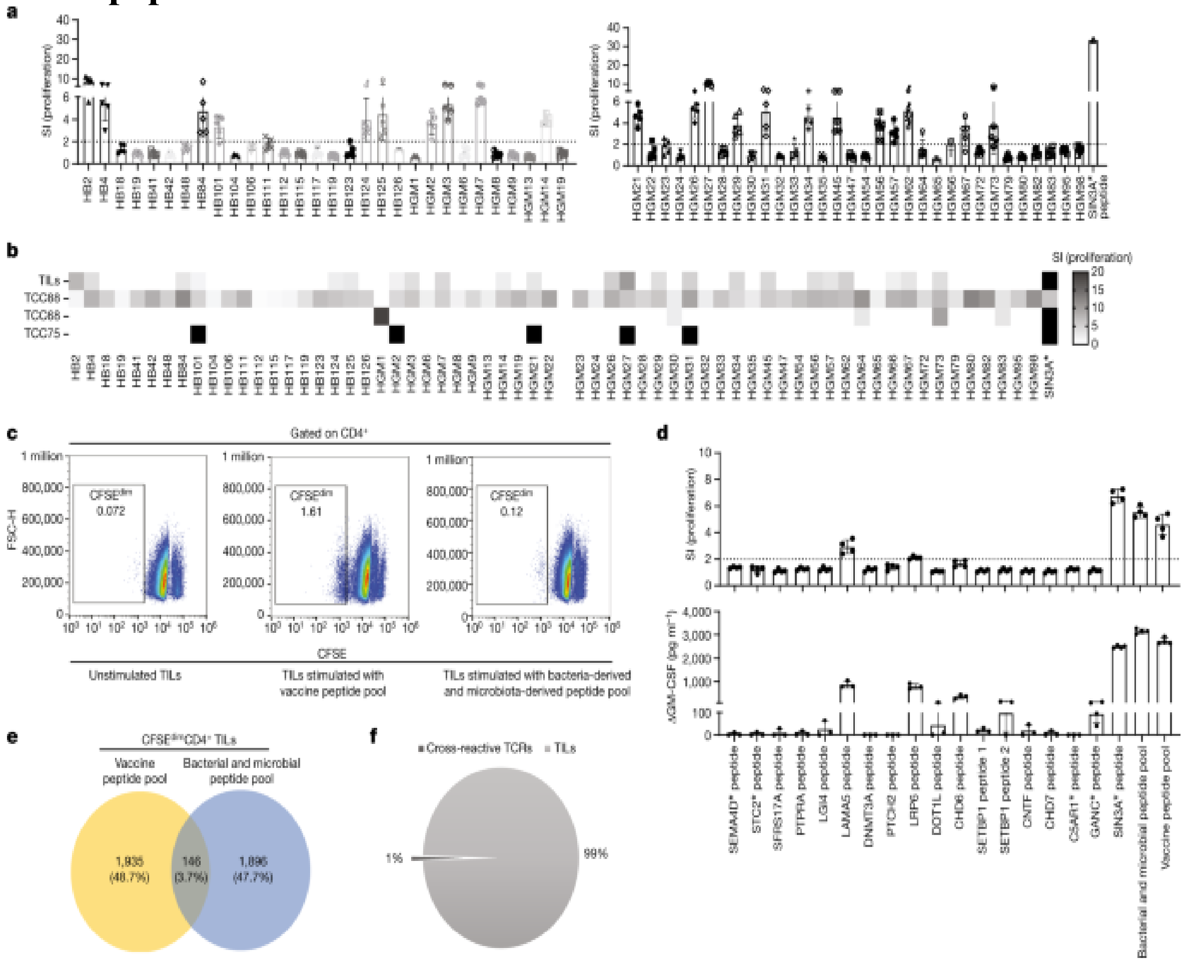
TILs recognize bacterial peptides

The above observation raised the question whether the reactivity of TCC88 against tumour antigens is unique for this clone and whether bacteria-derived and microbiota-derived peptides could also be recognized by bulk TILs. We therefore tested the proliferative and cytokine responses of bulk TILs against the bacteria-derived and microbiota-derived peptides that had been identified as targets for TCC88. Indeed, bulk TILs also recognized 23 bacteria-derived and microbiota-derived peptides (Fig. 4a and Extended Data Fig. 5c). To exclude that the response of bulk TILs is due to the fact that they contain TCC88, we tested additional TIL-derived TCCs that had

been established by stimulation with SIN3A* peptide against the bacteria-derived and microbiota-derived peptides. TCC68 and TCC75 also recognized both SIN3A* peptide and various bacteria-derived and microbiota-derived peptides (Extended Data Fig. 5d). However, not all SIN3A* peptide-responsive TCCs showed this pattern. As an example, TCC8F7 strongly responds to the SIN3A*-derived and several other glioblastoma-derived peptides, but not any of the bacterial antigens (Extended Data Fig. 5d). Figure 4b summarizes the responses of TILs and TCCs to bacteria-derived and microbiota-derived peptides. Their reactivity against the peptides that had been selected based on the unbiased antigen search for TCC88 differed, but both TCC68 and TCC75 recognized several bacteria-derived and microbiota-derived peptides with different strengths (Fig. 4b). The comparison of the reactivity of TCC88, bulk TILs and two other TCCs shows that TILs contain multiple TCCs with different reactivity patterns but clear responses to the bacteria-derived and microbiota-derived peptides. To corroborate this point, we compared the TCR expression of the TCCs. TCC88 expresses TCR beta variable 6.1 (TCRBV6.1), and TCC68 and TCC75 express TCRBV6.2. We further stained TILs using a panel of 22 fluorescently labelled TCRBV antibodies that can detect 27 different TCRBVs (Supplementary Table 12), excluding TCRBV6.1 and TCRBV6.2. Stained $CD4^+$ TILs were then sorted and rested for a day before testing them with peptides (Extended Data Fig. 6a). The antibody panel stained approximately half of the TCRVβs in the TILs and not only excluded TCRBV6.1 and TCRBV6.2 but also those for which no antibody was available. These depleted TILs still responded to bacterial peptides (shown with red arrows in Extended Data Fig. 6b), supporting that bulk TILs contain several T cells that recognize bacteria-derived and microbiota-derived peptides. Moreover, the cytokine and chemokine profile of TCC75 shows that it also secretes pro-inflammatory cytokines upon activation by either a tumour neoantigen or *Prevotella* spp. (HGM27 peptide). Of note, TCC75 secretes higher levels of IL-10 upon stimulation with tumour antigen, showing that different TCCs may secrete different cytokines and chemokines upon stimulation with a similar antigen (Extended Data Fig. 5e). Next, we compared the bacteria-derived and microbiota-derived peptides that we identified for TIL-derived TCC88 with the IPdBPs of the tumour, but did not find shared sequences. However, some peptides derived from the same source bacterium including *Bacillus* sp. and *Bacteroides ovatus*,

indicating that TCC88 responded to peptides from different proteins of these bacteria (Supplementary Table 15). These results demonstrate that unbiased epitope discovery for a single TIL-derived TCC was informative regarding the reactivity of bulk T cells, in that TILs were able to respond to various bacteria-derived and microbiota-derived peptides.

Fig. 4: Response of bulk TILs to bacteria-derived and microbiota-derived peptides.



a, Bulk TILs were stimulated with bacteria-derived and microbiota-derived peptides and irradiated autologous PBMCs as APCs. Proliferation was measured by ^{3}H-thymidine incorporation after 5 days (data represent the mean value of five replicate wells ± s.e.m.). **b**, Overview of proliferative responses of several TCCs and bulk TILs to bacteria-derived and microbiota-derived peptides shown in a heatmap. SI < 2 are shown in white and SI ≥ 2 are shown in varying shades of grey. **c**, Bulk TILs were labelled with CFSE and separately stimulated with a pool of bacteria-derived and

microbiota-derived peptides and a pool of vaccine or tumour peptides that includes peptides from somatic mutations and genes that were overexpressed in the autologous tumour (see Supplementary Table 13). Peptide-pulsed, irradiated autologous PBMCs were used as APCs, and bulk TILs were incubated for 10 days. Sixty replicate wells were pooled and proliferating $CD4^+$ T cells ($CFSE^{dim}$) were sorted. FSC-H, forward scatter height. **d**, Expanded $CFSE^{dim}CD4^+$ TILs that responded to the bacteria-derived and microbiota-derived peptide pool were stimulated with several tumour antigens, a bacteria-derived and microbiota-derived peptide pool and a pool of vaccine peptides. After 5 days, supernatants were harvested and proliferation was measured using ^{3}H-thymidine incorporation (data represent the mean SI value of four replicate wells ± s.e.m.). GM-CSF was measured by ELISA and is depicted in (pg ml^{-1}) after subtraction of the negative control. **e**, Part of the $CFSE^{dim}CD4^+$ TILs that responded to a bacteria-derived and microbiota-derived peptide pool and a vaccine peptide pool were used for TCRVβ sequencing. The percentage and frequency of TCR overlap of unique productive TCRVβ sequences of $CFSE^{dim}$ TILs responding to the vaccine peptide pool and TCRVβ sequences of $CFSE^{dim}$ TILs responding to the bacteria-derived and microbiota-derived peptide pool are shown. **f**, TCR sequence overlap of cross-reactive T cells in TILs recognizing that both the vaccine-derived peptide pool and the bacteria-derived and microbiota-derived peptide pool comprise 1% of entire TILs. Also see Extended Data Figs. 5 and 6.

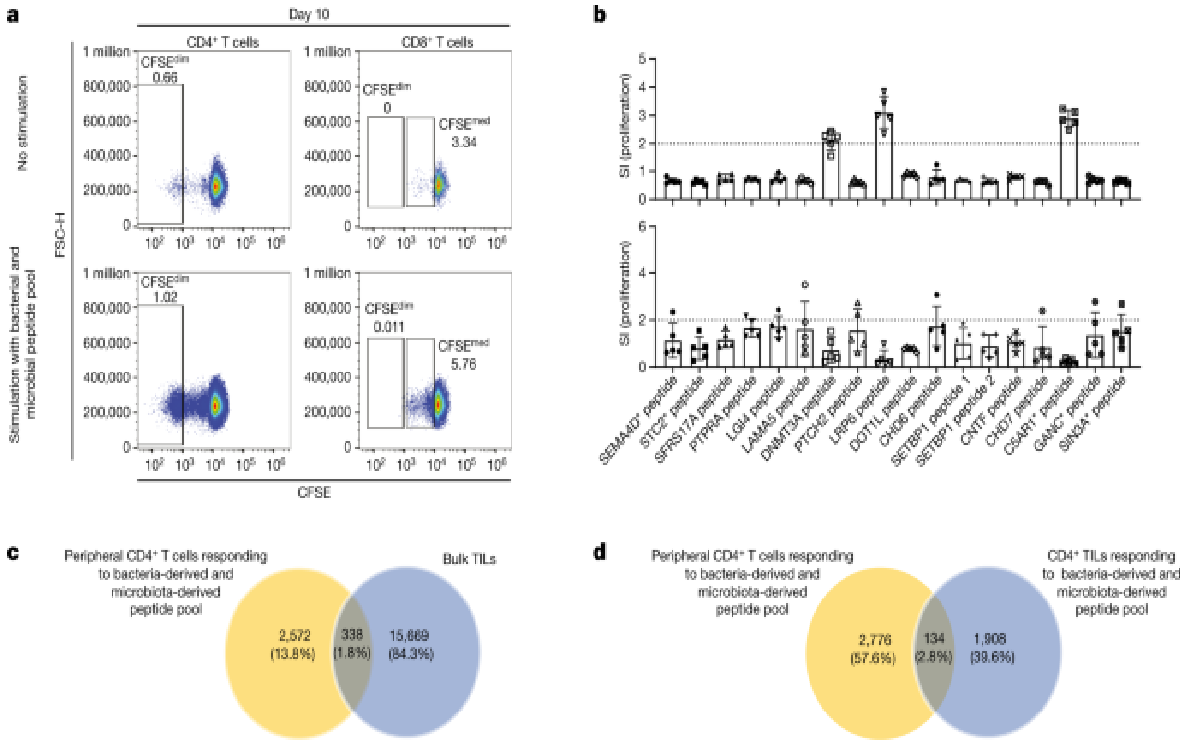
Source data

Cross-reactivity in bulk TILs

As several TIL-derived TCCs cross-recognize a tumour antigen and bacteria-derived and microbiota-derived peptides, we addressed whether bulk TILs responding to bacteria-derived and microbiota-derived peptides also recognize tumour antigens. CFSE-labelled TILs were stimulated with a pool of 22 peptides that had been applied in the vaccine of the patient and included somatic mutations and peptides derived from genes that were overexpressed in the autologous tumour[25] (Fig. 4c and Supplementary Table 13). CFSE-labelled TILs were also separately stimulated with a pool of 23

bacteria-derived and microbiota-derived peptides that were recognized by bulk TILs (Fig. 4c and Supplementary Table 13). Proliferative ($CFSE^{dim}$) $CD4^+$ T cells were sorted, and part of these were short-term expanded with phytohaemagglutinin (PHA) and IL-2 and the remainder used for TCR sequencing (Fig. 4c). A higher number of T cells within the bulk TILs responded to the vaccine peptide pool than to the bacteria-derived and microbiota-derived peptide pool (Fig. 4c). However, $CD4^+$ TILs that had been stimulated with the bacteria-derived and microbiota-derived peptide pool also recognized several tumour antigens including laminin subunit α5 (LAMA5), low-density lipoprotein receptor-related protein 6 (LRP6) and SIN3A* peptides, as well as the pools of tumour antigens and bacteria-derived and microbiota-derived peptides by proliferation and/or GM-CSF secretion (Fig. 4d). Moreover, short-term expanded $CD4^+$ TILs that responded to the bacteria-derived and microbiota-derived peptide pool showed a comparable cytokine and chemokine profile including several pro-inflammatory cytokines (IFNγ, TNF and granzyme B) after activation with both peptide pools (Extended Data Fig. 6c, left panel). When we compared TCR sequences of the $CFSE^{dim}CD4^+$ T cells responding to the tumour vaccine pool with those responding to the bacteria-derived and microbiota-derived peptide pool, they share 3.7% of the TCRs, indicating that these T cells are cross-reactive against tumour and bacterial antigens (Fig. 4e). The fraction of TCRs that are found in both expanded T cell fractions makes up 1% of the entire TILs (Fig. 4f). From the above TIL-derived TCCs, we only identified TCC75 and TCC88 in the TCR data of both $CFSE^{dim}$ fractions, indicating that the fractions of cells that were isolated for TCR sequencing were too small.

Fig. 5: Peripheral memory T cells recognize several tumour antigens after activation with bacteria-derived and microbiota-derived peptides.



a, CD45RA$^-$ peripheral T cells were isolated and labelled with CFSE. Cells were then stimulated with the bacteria-derived and microbiota-derived peptide pool for 10 days. Next, 60 replicate wells were pooled and proliferative (CFSE^{dim}) CD4$^+$ T cells were sorted and short-term expanded. **b**, Bacteria-reactive or microbiota-reactive peripheral CD4$^+$ T cells responded to several tumour antigens from the autologous tumour when stimulated with peptide presented by irradiated autologous PBMCs as APCs (top). Proliferation was measured after 5 days by ^{3}H-thymidine incorporation (data represent the means of five replicate wells ± s.e.m.). Peripheral blood memory T cells were isolated and directly tested with tumour antigens (bottom). Proliferation was measured after 5 days by ^{3}H-thymidine incorporation (data represent the means of five replicate wells ± s.e.m.). **c**,**d**, Part of CFSE^{dim} (that is, proliferating) peripheral CD4$^+$ T cells that responded to the bacteria-derived and microbiota-derived peptide pool was used for TCRVβ sequencing. The percentage and frequency of TCR overlap of unique productive TCRVβ sequences of bacteria-reactive and microbiota-reactive peripheral CD4$^+$ (CFSE^{dim}) and TCRVβ sequences of bulk TILs (**c**), and the overlap of unique productive TCRVβ sequences of bacteria-reactive and microbiota-reactive peripheral CD4$^+$ (CFSE^{dim}) and

TCRVβ sequences of (CFSE^{dim}) bacteria-reactive and microbiota-reactive TILs (**d**) are shown. Also see Extended Data Figs. 6 and 7.

Source data

Stimulation of peripheral T cells

When considering bacteria-derived and microbiota-derived peptides for future vaccination approaches, it will be important to show that they induce tumour-specific immune responses in the peripheral immune system. To address whether T cells with reactivity against bacterial peptides from the TCC88 profiling also stimulate peripheral blood memory cells, we CFSE-labelled CD45RA^{-} cells from peripheral blood mononuclear cells and activated these with the above pool of bacteria-derived and microbiota-derived peptides that were stimulatory for TILs. CD4^{+} T cells responded substantially stronger than CD8^{+} T cells to this peptide pool (Fig. 5a). Proliferating CFSE^{dim}CD4^{+} T cells were then sorted and short-term expanded with PHA and IL-2 (Fig. 5a). CFSE^{mid}CD8^{+} T cells were not tested further.

We next addressed whether the sorted and short-term expanded peripheral memory CD4^{+} T cells, which had been pre-selected by a pool of bacteria-derived and microbiota-derived peptides, responded to the peptides derived from tumour-specific antigens and tumour-associated antigens of the autologous tumour. These cells recognized three tumour antigens via proliferation, whereas unselected peripheral memory T cells did not (Fig. 5b). In addition to antigen-specific proliferation, they produced various cytokines and chemokines in response to the LRP6 peptide in the selected population compared with the unselected cells (Extended Data Fig. 6c, right panel). To address whether these peripheral blood CD4^{+} T cells can be found in TILs, we compared the TCR sequences of bacteria-derived and microbiota-reactive peripheral CD4^{+} memory T cells with bulk TILs (Fig. 5c) and TILs selected by stimulation with the bacteria-derived and microbiota-derived peptide pool (Fig. 5d). Of the peripheral blood-derived bacteria-responsive and microbiota-responsive T cells, 1.8% are also present in bulk TILs (Fig. 5c). Furthermore, 2.8% of the T cells are present in both

the selected peripheral T cells and TILs responding to the bacteria-derived and microbiota-derived peptide pool (Fig. 5d).

To return to our starting observation with IPdBPs, we examined whether T cells with specificity for these also exist in peripheral blood memory T cells and whether they respond to autologous tumour antigens. We stimulated CFSE-labelled peripheral memory $CD4^+$ T cells with all 37 IPdBPs, sorted the proliferating cells and short-term expanded them (Extended Data Fig. 7a). These bacteria-activated memory T cells recognized the same three tumour antigens, which are recognized by peripheral T cells that had been selected with bacteria-derived and microbiota-derived peptides via proliferation and/or GM-CSF cytokine expression as indicators for T cell activation (Extended Data Fig. 7b). Furthermore, the cytokine and chemokine profile of immunopeptidome-reactive peripheral $CD4^+$ T cells stimulated with the tumour peptide (LRP6) as well as the pool of IPdBPs showed that T cells secreted not only GM-CSF but also several other pro-inflammatory cytokines including IFNγ, TNF and granzyme B (Extended Data Fig. 7c). To address whether there was an overlap between peripheral blood memory T cells responding to IPdBPs and the bulk TILs, we sequenced the TCRVβ repertoires of both populations and found that indeed 0.5% of the TCR sequences were shared (Extended Data Fig. 7d). This indicates that T cells responding to bacterial antigens, which are presented in the tumour, can also be found in the peripheral immune compartment. As another means to show that peripheral blood $CD4^+$ T cells that had been activated by bacteria and microbiota or IPdBPs recognized the same tumour antigens, we compared the TCR sequences of these two populations. These data support the cross-reactivity results. Of the TCR sequences, 3.4% were shared between the two populations, and most of these TCRs were clonally expanded in each group (Extended Data Fig. 7e). Finally, to assess whether the bacteria-derived and microbiota-derived peptides that were recognized by TCC88 and TILs from patient 1635WI were only stimulatory for T cells of this patient or also for TILs from another patient with glioblastoma, we tested TIL reactivity from another patient with glioblastoma (GBM-40), and indeed these TILs recognized several bacteria-derived and microbiota-derived peptides from patient 1635WI, albeit less strongly (Extended Data Fig. 8a,b).

These data show that (1) T cells with reactivity against bacterial peptides that are presented on HLA class II molecules, that is, the immunopeptidome, within the tumour exist in the peripheral blood, (2) memory T cells that recognize bacteria and microbiota that had been identified by analysing a tumour neoantigen-reactive TCC can be found in the periphery and TILs, and (3) a fraction of T cells from both (1) and (2) also responds to tumour antigens (Extended Data Fig. 9a,b).

Discussion

Our study examines the role of bacteria for tumour antigen-specific immune responses in glioblastoma, a tumour with poor prognosis and high unmet medical need[36]. We build on recent observations about the involvement of bacteria and particularly gut microbiota and how they may shape immune responses[1,2,37,38,39]. Findings of bacterial communities in tumours including glioblastoma[20], bacterial peptides in the immunopeptidome of tumours[21] and the influence of commensal bacteria and a bacteriophage on stimulating tumour-specific immune responses[12,13] motivated our studies. We first confirmed data by Nejman et al.[20] that bacterial sequences are found in glioblastoma and by Kalaora et al.[21] regarding bacteria-derived and microbiota-derived peptides in the HLA class II immunopeptidomes of patients with melanoma[21], but here in glioblastoma. Peptides and bacterial strains, from which IPdBPs derived, varied between 5 and more than 50. Hence, compared with tumour-derived human peptide sequences, which usually are a few thousands[24] (Fig. 1c, top), the abundance of bacterial peptides in the HLA class II immunopeptidomes was low. This may be due to low bacterial burden in tumour cells[40] and/or on tumour-infiltrating APCs where bacterial peptides may occupy only a minority of HLA class II molecules and fall below the detection limit[21,41]. Consistent with Nejman et al.[20], we found the phyla Proteobacteria and Firmicutes to be most prevalent in glioblastoma. Although we first focused on IPdBPs, both analyses identified bacterial antigens from Enterobacterales, Burkholderiales, Bacillales, Clostridiales and others at the order-level phylotypes[20]. In a patient, whom we had previously treated with a personalized vaccination using tumour neoepitope peptide cocktails, we tested whether IPdBPs are recognized by TILs. Compared with the mutated SIN3A* peptide, one of the

vaccine peptides, TILs responded very weakly to 3 out of 37 IPdBPs via proliferation; however, secretion of GM-CSF and TNF as well as CXCL10, which enhances cytotoxic $CD8^+$ T cell trafficking to tumours[42], support the reactivity. CXCL10 has been used to predict responses to anti-PD1 treatment in several tumours[43]. When examining more pro-inflammatory SIN3A*-reactive TCCs, two out of ten recognized single (TCC57) or multiple (TCC2D3) IPdBPs, demonstrating that T cells with cross-reactivity against tumours and IPdBPs can be found in the TILs. With an unbiased antigen discovery approach, we had previously identified targets for autoantigen-specific, virus-specific and bacteria-specific $CD4^+$ TCCs, and all of these recognized multiple antigens[26,27,28,29,32]. Therefore, we addressed, now starting from tumour-infiltrating T cells, whether TIL-derived TCCs respond to viral-derived, bacterial-derived, microbiome-derived, human protein-derived or tumour antigen-derived peptides. TIL-derived, SIN3A*-specific TCC88 served as a prototypic example to search for target antigens among sequences of all human proteins, viruses and bacteria with tropism for humans, gut microbiota, the translated proteome of the autologous primary and recurrent tumours and the tumour immunopeptidome. The sizes of these databases ranged from 16,000 (immunopeptidome) to more than 3.4 billion (human gut microbiota) decamer peptides (Fig. 3a). We synthesized 542 peptides from different sources with predicted high stimulatory scores for TCC88 (Fig. 3a). TCC88 responded to 109 peptides, and, among these, reactivity was broadest to bacterial and microbiota-derived peptides, but also to many tumour proteome-derived peptides (Fig. 3a). Of 226 bacterial and microbiota peptides, TCC88 recognized 63, from which 34 stimulated TCC88 equally well or better than the cognate antigen. Besides *B. ovatus* and *Bacillus* sp., other bacteria-derived and microbiota-derived peptides were derived from completely different bacterial species than IPdBPs (Extended Data Fig. 9c and Supplementary Table 15). These data show that tumour-infiltrating TCCs may cross-recognize not only single bacterial, viral or phage-derived peptides[12,13,44,45] but also large numbers of peptides from both pathogenic bacteria and gut microbiota-derived commensals. Similar to previous data with autoreactive TCCs in MS, we did not identify amino acid patterns or HLA-binding or TCR-binding motifs that allow to distinguish stimulatory from non-stimulatory peptides[46] (Extended Data Fig. 5b) to explain the broad cross-reactivity[47]. The primary HLA restriction element of TCC88 is DRB3*02:02, and similar to virus-specific and autoreactive TCCs

that we had previously described[31,32], TCC88 is cross-restricted, that is, capable of recognizing peptides also in the context of other autologous HLA-DR molecules[31,48]. Regarding its function, TCC88 showed in vitro cytotoxicity against an autologous glioblastoma cell line and a multifunctional pro-inflammatory phenotype with strong secretion of TNF, IFNγ, granzyme B, GM-CSF and others. Although the response to SIN3A* and two bacterial peptides is almost identical, the relatively higher secretion of IL-10 upon stimulation with the bacterial peptides may indicate that the latter could both activate, but also dampen, intratumoural T cell responses.

These data raised important questions including whether TCC88 is an unusually cross-reactive clone among the TILs and whether the bacteria-derived and microbiota-derived peptides recognized by TCC88 could also stimulate bulk TILs. Indeed, both were the case. Of the 63 bacteria-derived and microbiota-derived peptides that were recognized by TCC88, 23 also stimulated bulk TILs, and several of these were also recognized by other SIN3A*-specific, TIL-derived TCCs, underscoring the broad reactivity of $CD4^+$ TCC against bacterial and microbial peptides. Furthermore, the diversity of the TIL response against bacteria and microbiota is supported by the fact that there are SIN3A*-specific TCCs (TCC68 and TCC75) that responded to bacteria-derived and microbiota-derived peptides that had been identified as targets for TCC88, whereas other clones such as TCC8F7 did not recognize them (Extended Data Fig. 9a).

Hence glioblastoma-infiltrating TILs can recognize multiple peptides from bacteria and gut microbiota, and the immunopeptidome analyses reveal that such peptides might be locally presented in the tumour (Extended Data Fig. 9a). If one wanted to use this knowledge for tumour vaccination in the future, it is of interest to examine whether tumour-derived, bacteria-derived and microbiota-derived peptides also stimulate peripheral blood memory T cells. Indeed, upon activation of peripheral blood memory T cells with bacteria-derived and microbiota-derived peptides derived either from the target search for TCC88 or from the tumour immunopeptidome not only identified reactivity against bacterial targets from both sources but also against tumour antigens (Extended Data Fig. 9a). Moreover, comparing the TCR repertoires of bulk peripheral blood memory T cells that had been pre-stimulated with either bacteria-derived and microbiota-derived peptides or

with several tumour-derived antigens implies that cross-reactivity, which was shown at the clonal level for TCC88, probably exists for more T cells within these T cell populations (Extended Data Fig. 9b). Of note, one tumour antigen, a peptide from LRP6, was recognized by bulk T cells after pre-stimulation with both bacterial and microbial peptide and IPdBP peptide pools. LRP6 is a co-receptor in the WNT signalling pathway that supports the proliferation of tumour cells, and several lines of evidence have shown that inhibition of LRP6 leads to suppression of tumour cell proliferation and enhances chemosensitivity of cancer cells[49]. Hence, the reactivity of peripheral blood memory T cells overlaps with that of TILs.

Our data highlight several important points. The immunopeptidome analyses suggest that bacteria and gut microbiota can exist within glioblastoma and that peptides from these are presented by HLA molecules on tumour cells and probably also on local APCs. Thus, we confirm previous data and extend these findings by showing that some IPdBPs are weakly recognized by both individual tumour-specific TCCs and bulk $CD4^+$ TILs. Furthermore, systematic unbiased epitope discovery of tumour-specific TCCs identified novel bacteria-derived and microbiota-derived antigens, which were highly stimulatory for clones and bulk TILs. Remarkably, we confirm that individual TCCs recognize not only single bacterial or phage-derived peptides[13,44] but are also broadly cross-reactive with multiple bacteria-derived and microbiota-derived peptides. Such cross-reactivity patterns may explain the response of patients with cancer with certain gut microbiota composition to immune checkpoint inhibitors. Vétizou and colleagues found that *B. fragilis* induced a T_H1 phenotype, which led to restoration of treatment responsiveness to CTLA4 blockade in tumour-bearing germ-free mice. Mice transplanted with faeces of CTLA4-responsive patients with melanoma with high *B. fragilis* abundance also responded to the blockade[2]. Moreover, *Ruminococcus bromii* is prevalent in the gut of patients with melanoma who respond to anti-PD1 therapy[38], and *Faecalibacterium* of the Ruminococcaceae family was proposed to favour such responses[38]. Other bacteria such as *Alistipes* spp. and *Eubacterium* spp. are enriched in the gut microbiota of patients with non-small-cell lung carcinoma responding to anti-PD1 therapy[1]. *Alistipes indistinctus* has been shown to restore the response to checkpoint blockade in mice[1]. Using our unbiased approach, we found peptides from *B. fragilis*, *R. bromii*, *Alistipes* and *Eubacterium* genus

that stimulated TCC88 (Supplementary Table 15). This observation provides further hints at T cell cross-reactivity as a possible mechanism of effective immunotherapies due to certain gut microbiota composition. Our data indicate that not only commensals but also peptides from certain pathogenic bacteria may be exploited for strong stimulation of T cells against cancer cells (Supplementary Table 15). The reactivity of peripheral blood memory T cells against both bacteria-derived and microbiota-derived as well as tumour peptides may be used for vaccination approaches that combine tumour-derived and bacterial and microbial peptides, use mRNAs encoding such combinations of tumour/self and foreign antigens or even adoptive T cell therapies with TCCs such as TCC88 or TCC2D3 or their TCRs. The recent observation that bacteria can be involved in tumour metastasis[40] suggests that T cell activation against bacteria and tumour cells may not only be used early in the disease but even at later stages after cancer metastasis.

Limitations of our study include that we could not examine the ‘autologous’ gut microbiota of patients, their lung microbiome, which appears to have an important role in brain homing of pro-inflammatory T cells[50], and also that the most detailed studies on both tumour immunopeptidome and its annotation to bacteria and microbiota, unbiased target discovery for TIL-derived TCC and their cross-reactivity against bacteria and tumour antigens mainly derive from one patient. However, we identified TIL reactivity against some of the bacteria-derived and microbiota-derived peptide targets from patient 1635WI already in TILs from another patient with glioblastoma (Extended Data Fig. 8). Furthermore, the three different methodologies, that is, 16S rRNA sequencing, immunopeptidomics and unbiased target discovery for a TIL-derived T cell clone, identify largely non-overlapping bacterial and gut microbiota-derived species and peptides. All suggest an involvement of bacteria and gut microbiota in the biology of the tumour and/or immune defence against it; however, future studies will need to examine their roles in more depth.

To explore the above therapeutic applications further, strong emphasis would have to be laid on rapid analysis of the tumour immunopeptidome for bacterial and microbial sequences and/or, even better, the unbiased target discovery for TIL-derived clones, and de-risking potential reactivity against vital organs should be incorporated. Our findings support that this is, in principle, possible and that intratumoural bacteria and gut microbiome

communities probably have a role in tumour defence mechanisms and could be exploited systematically in the context of immune checkpoint inhibitors and personalized cancer therapies[1,2,37,38,39].

Methods

Patient information and materials

Patient 1635WI, who was treated with a personalized neopeptide vaccination, was a 55-year-old male diagnosed with glioblastoma. After resection, he was treated with standard radiochemotherapy and additionally with the immune checkpoint inhibitor pembrolizumab (anti-PD1) and the anti-VEGF monoclonal antibody bevacizumab. He was interested in participating in an individual treatment attempt with a personalized peptide vaccination therapy[25]. He was fully informed about the potential risks of such a vaccination, which, by definition, is an individual medical treatment ('compassionate use') and not subject to special regulations for medical research according to the Therapeutic Products Act (TPA) or Swiss Federal Human Research Act (HRA). Consequently, individual medical treatments are not approved by the Ethics Committee or Swiss Agency for Therapeutic Products (Swissmedic), as they do not constitute clinical trials for the purpose of systematically gaining knowledge. The patient gave written informed consent to receive the individualized vaccination.

The cohort of 18 additional patients with glioblastoma includes the primary tumour of 16 patients and 2 recurrent tumours from 2 patients, all from the University Hospital Zurich, Zurich, Switzerland. Tumour tissues were immediately frozen after surgery and stored at −80 °C until immunoaffinity purification of HLA–peptide complexes. Autologous tumour cells were also isolated and frozen from several samples ($n = 6$) and immunopeptidomes were analysed (Supplementary Table 1). Written informed consent was obtained from each patient in accordance with the local ethical requirements (KEK-ZH-Nr. 2015-0163). All included patients underwent elective brain tumour surgery at the Department of Neurosurgery at the University Hospital Zurich and were diagnosed with glioblastoma (IDH1 wild type) by a board-certified neuropathologist. No other inclusion or exclusion criteria were applied so we do not expect other biases. Sample size was determined

based on the availability of biological materials. To generate autologous tumour cell lines, tumour tissue was washed thoroughly using PBS. Next, they were cut into small pieces and treated with digestion mix including RPMI-1640 medium (Sigma-Aldrich) containing 10% FCS (Eurobio), 1 mg ml^{-1} of collagenase IV (Sigma-Aldrich) and 0.1 mg ml−1 of DNase I (Sigma-Aldrich), and incubated for 42 min at 37 °C. The reaction was stopped by adding PBS with 2 mM EDTA (Invitrogen) and the lysate was passed through a filter mesh. Next, tumour cells were isolated using Percoll (Sigma-Aldrich) gradient. Isolated cancer cells were passaged up to 30 times to make sure no other cells would survive before they were snap frozen and used for the immunopeptidome analyses. All cells were regularly tested for mycoplasma contamination and were negative.

DNA extraction and genotyping

DNA was extracted from PBMCs, tumour cells and tissues of patients with glioblastoma using DNeasy Blood & Tissue Kits (Qiagen). The genotyping of HLA class I and HLA class II was done by high-resolution HLA sequence-based typing (Histogenetics) (Supplementary Table 2).

PBMCs, cell lines and TCCs

PBMCs of patient 1635WI were isolated from a leukapheresis, and allogenic PBMCs were isolated from buffy coats by Ficoll (Eurobio) density gradient centrifugation.

BLS-DRB1*03:01 (kindly gifted by W. W. Kwok, Benaroya Institute, WA, USA), BLS-DRB1*04:02, BLS-DRB3*02:02 and BLS-DRB4*01:01 are EBV-infected B cells from a patient with BLS type II that are transfected with single HLA-DR molecules of the patient with glioblastoma. Thus, each BLS-DR transfectant expresses a single HLA-DR molecule. In brief, various HLA-DRB complementary DNAs were inserted after the CMV promoter in a pLenti-CMV vector (kindly provided by M. Scharl, University of Zurich). Plasmid vectors carrying different HLA-DRBs were used to prepare lentiviruses using pCMV-dR8.91 (packaging plasmid) and pMD2.G (VSV-G envelope plasmid) (both kindly gifted by M. Scharl). Next, BLS cells expressing HLA-DRA1*01:01 (kindly provided by W. W. Kwok) were

transduced with various HLA-DRB lentiviruses and cells were selected using blasticidine after 7 days. BLS cells were maintained and expanded using RPMI-1640 medium (Sigma-Aldrich) containing 10% heat-inactivated FCS (Eurobio), 2 mM l-glutamine (Thermo Fisher Scientific), 100 U ml^{-1} penicillin (Corning) and 100 μg ml^{-1} streptomycin (Corning). BLS cell lines were regularly examined for expression of HLA-DR via FACS. Bulk TILs and isolated TCCs including TCC88 and other TCCs were expanded using irradiated allogenic PBMCs (45 Gy) in T cell medium containing IMDM (GE Healthcare), 5% heat-inactivated human serum (Blood Bank Basel, Switzerland), 1 μg ml^{-1} PHA, 20 U ml^{-1} of human IL-2 (h-IL-2), 2 mM l-glutamine (Thermo Fisher Scientific), 100 U ml^{-1} penicillin (Corning) and 100 μg ml^{-1} streptomycin (Corning). The medium was changed every 3–4 days and new h-IL-2 was added each time.

Proliferation assay

All peptides were synthesized with acetylated N terminus and amidated C terminus (Peptides & Elephants) (Supplementary Table 10). Proliferation of TCCs and bulk TILs was measured using the ^{3}H-thymidine incorporation (Hartmann Analytic) assay after 3 and 5 days, respectively. Isolated TCCs (2×10^4 cells per well) were co-cultured with irradiated BLS cells (300 Gy) (5×10^4 cells per well) in 200 μl of X-VIVO 15 medium (Lonza) and seeded in 96-well U-bottom plates (Greiner Bio-One). Bulk TILs (5×10^4 cells per well) were co-cultured with irradiated autologous PBMCs (45 Gy) (2×10^5 cells per well) in 200 μl of X-VIVO 15 medium (Lonza) and seeded in 96-well U-bottom plates (Greiner Bio-One). Peptides were then used at a final concentration of 10 μM to stimulate the TCCs or bulk TILs. Results are shown as SI, that is, the ratio of count per minute in the presence of the peptides divided by the negative control (no peptide). SI ≥ 2 is considered positive.

For testing peripheral blood memory T cells, $CD45RA^{-}$ cells were isolated using human CD45RA microbeads (Miltenyi). Negatively selected $CD45RA^{-}$ cells were seeded (2×10^5 cells per well) in 200 μl of X-VIVO 15 medium (Lonza) in 96-well U-bottom plates (Greiner Bio-One). Peptides

were then added at a final concentration of 10 µM and incubated for 7 days before measuring their proliferation via ^{3}H-thymidine incorporation.

To isolate antigen-reactive T cells from the TILs or peripheral T cells, CFSE (Sigma-Aldrich) was used to label the cells. In brief, cells were incubated with CFSE (final concentration of 0.5 µM) for 3 min at room temperature. Next, the labelling was immediately stopped using a serum-containing medium and washed before cells were seeded (5×10^4 cells per well) in the presence of irradiated autologous PBMCs (45 Gy) (2×10^5 cells per well) in 200 µl of X-VIVO 15 medium (Lonza) and the peptide of interest. After 10 days, cells were pooled, washed once with X-VIVO 15 medium (Lonza) and stained with live-dead yellow (Invitrogen) as well as surface markers and cells sorted with a SH800S Cell Sorter (Sony).

Sorting bulk T cells

TILs expressing specific TCRBVs were isolated from bulk TILs using a panel of 22 fluorescently labelled TCRBV antibodies that detect 27 different TCRBVs. Twelve TCRBV antibodies were labelled with PE and 10 with FITC (Beckman Coulter). Five million TILs that had been expanded once with the above protocol were thawed and washed once with PBS before incubation with human IgG (1:25) (Sigma-Aldrich) for 30 min to block Fc fragments and reduce nonspecific antibody staining. After another wash, cells were stained with CD4 (1:100) (APC anti-human CD4, eBioscience), CD8 (1:100) (Pacific blue anti-human CD8, BioLegend), TCRBV-PE (1:25) (Beckman Coulter) and TCRBV-FITC (1:25) (Beckman Coulter) antibodies (Supplementary Table 12) for 30 min. Finally, cells were sorted, seeded at 5×10^4 cells per well in T cell medium without h-IL-2 and rested for 1 day before stimulation with 2×10^5 cells per well of irradiated autologous PBMCs (45 Gy) and peptide (10 µM).

Furthermore, TILs were labelled with CFSE and 5×10^4 cells per well were stimulated with irradiated autologous PBMCs (2×10^5 cells per well) and pools of 22 glioblastoma-associated peptides that had been used in an individualized vaccination of patient 1635WI[25] or 23 bacteria-derived and microbiota-derived peptides, which had been identified by an unbiased target search for TCC88 (Supplementary Table 13) (see below). Both

peptide pools have a final concentration of approximately 25 μM for all the peptides in the pool, and the DMSO concentration was always lower than 0.5%.

Regarding isolation of peripheral memory T cells, $CD45RA^{-}$ cells were sorted using human CD45RA microbeads (Miltenyi) and labelled with CFSE. Next, 2×10^5 cells per well were seeded in 200 μl of X-VIVO 15 medium (Lonza) in 96-well U-bottom plates (Greiner Bio-One) in the presence of the bacteria-derived and microbiota-derived peptide pool or the IPdBP pool. The IPdBP pool contains 37 HLA class II-derived bacterial peptides at a final concentration of 20 μM for all the peptides in the pool (Supplementary Table 13). After 10 days, cells were Fc-blocked using human IgG (1:25) (Sigma-Aldrich) and labelled with live-dead yellow (1:1,000) (Invitrogen), PerCP-Cy5.5 anti-human CD3 (1:100) (BioLegend), APC anti-human CD45RA (1:50) (BioLegend), APC-Cy7 anti-human CD4 (1:100) (BioLegend) and Pacific blue anti-human CD8 (1:100) (BioLegend). $CD45RA^{-}CD3^{+}CD4^{+}CFSE^{dim}$ cells were then sorted and 2×10^4 cells per well were seeded in 96-well U-bottom plates (Greiner Bio-One) with irradiated allogenic PBMCs, PHA and h-IL-2 as described above. The medium was changed every 3–4 days and new h-IL-2 was added each time. If necessary, each well was split into two. After 14 days, cells were cryopreserved.

T cell cloning and characterization

To generate $CD4^{+}$ TCCs, bulk TILs were labelled with CFSE and stimulated with SIN3A* peptide (10 μM). After 10 days, live proliferative ($CFSE^{dim}$), $CD3^{+}$ (1:100) (PerCP/Cy5.5 anti-human CD3 antibody, BioLegend) and $CD4^{+}$ (1:100) (APC-Cy7 anti-human CD4 antibody, BioLegend) T cells were single-cell sorted into 96-well U-bottom plates (Greiner Bio-One). Cells were then expanded using irradiated allogenic PBMCs in the presence of PHA and h-IL-2 in T cell medium. The medium was changed every 3–4 days and h-IL-2 was added each time until day 12. Cells were then subjected to rapid expansion to expand TCCs. TCCs (5×10^4 cells) were transferred into T25 flasks (TPP) and were expanded using 3×10^7 irradiated allogenic PBMCs (45 Gy) in T cell medium containing anti-CD3 monoclonal antibody (OKT3) (Ortho Biotech). After 1 day, h-IL-2 was added at 50 U ml^{-1}; at day

5, the content of the flask was harvested, spun down and resuspended in new T cell medium containing 50 U ml^{-1} of h-IL-2. At day 8, cells were split into new T25 flasks and 20 U ml^{-1} of h-IL-2 was added. At day 16, cells were harvested, washed and cryopreserved in freezing medium including 90% heat-inactivated FCS and 10% DMSO (Applichem).

TCCs were characterized using primers for TCRVβ chains and further sequencing of amplicons[31]. In addition, the LEGENDplex Multi-Analyte Flow Assay kit (BioLegend) was used to identify the functional phenotype of TCC by measuring multiple cytokines in the supernatant following activation with anti-CD2/CD3/CD28 antibody-loaded MACSibead (Miltenyi). The HLA-DR restriction of TCCs was defined by co-culturing TCCs with the four aforementioned irradiated HLA-DR transfected BLS cells (300 Gy), presenting different concentrations of the cognate antigen and measuring proliferation after 3 days using ^{3}H-thymidine incorporation.

Cytokine measurements

Supernatants from each experiment were harvested a day before adding thymidine and frozen at −20 °C. Cytokines were measured using ELISA for human IFNγ ELISA MAX (BioLegend), human GM-CSF DuoSet ELISA (R&D Systems) and bead-based assays. A Synergy H1 hybrid multi-mode reader (Agilent) was used to measure the absorbance. The LEGENDplex Multi-analyte Flow Assay kit (BioLegend) is a 13-plex human T_H cytokine panel, which includes IFNγ, TNF, IL-2, IL-4, IL-5, IL-6, IL-9, IL-10, IL-13, IL-17A, IL-17F, IL-21 and IL-22. Staining was conducted as suggested by the manufacturer. LSR Fortessa Flow Cytometer (BD Biosciences) was used to measure cytokines, and data were analysed using FlowJo (Tree Star). The Human XL Cytokine Luminex Performance Assay Fixed Panel (R&D Systems) contains 45 analytes covering a broad range of various chemokines and cytokines. Bio-Plex 200 systems (Bio-Rad) was used to measure cytokine and chemokine secretion as suggested by the manufacturer.

ps-SCL and biometrical analysis

A library of l-amino acid decapeptide ps-SCL mixtures was prepared as mentioned before[30]. TCC88 is a $CD4^{+}$ TCC isolated from the TILs and is

restricted by HLA-DRB3*02:02. Irradiated BLS-DRB3*02:02 cells (300 Gy) were incubated with 200 mixtures from the library at 200 and 100 µg ml^{-1} concentration. TCC88 was then co-cultured with these cells. Each well contained 2×10^4 TCC88 and 5×10^4 BLS cells. After 72 h, supernatants were harvested and tested for IFNγ secretion by ELISA (BioLegend). Next, a scoring matrix was generated via $\log_{10}$ transformation of median IFNγ values of three independent experiments[26]. Next, biometrical analysis[29] was used to score all 10-mers from several large databases, that is, all human proteins (Swiss-Prot-20201202), the translated transcriptomes of the primary and recurrent autologous glioblastoma tumour (patient 1635WI), all viruses, which are annotated for infecting humans in UniProt (Swiss-Prot-20210305), as well as all bacteria (Swiss-Prot-20201202) and human gut microbiota[34]. To assess the validity of the in silico predictions of stimulatory peptides for TCC88, consecutive 10-mer peptides with the highest predicted scores in each database were synthesized and tested for their capacity to stimulate TCC88. Peptides were selected based on their predicted stimulatory scores and not other factors such as their source. Peptide scores ranged between 80% and 100% of the maximum theoretical score for all databases beside the tumour immunopeptidome (annotated by human proteins), which contained only a few peptides with a score over 80%. Therefore, we also synthesized several peptides within an extended score range between 60% and 80%. Peptides were tested at a final concentration of 10 µM on irradiated BLS-DRB3*02:02 and co-cultured with TCC88 as mentioned before. ^{3}H-thymidine incorporation (Hartmann Analytic) was used to measure proliferation.

T cell cytotoxicity assay

In brief, TCC88 (2×10^4 cells per well) was stimulated using the tumour antigen SIN3A* and *Victivallis* gut bacterium HGM3 peptides (2 µM) separately on irradiated autologous PBMCs (1×10^5 cells per well) for 72 h in X-VIVO 15 medium (Lonza). Next, the cells were pooled and washed once with X-VIVO 15 medium (Lonza). Autologous glioblastoma cells were stained with PKH26 Red Fluorescent Cell Linker Kit (Sigma-Aldrich) for 3 min and immediately washed three times with serum-containing medium. Target cells were then resuspended in X-VIVO 15 medium (Lonza) and

seeded at 2×10^4 cells per well. Activated and rested TCC88 cells were co-cultured at 1:1 and 5:1 effector to target ratios in duplicates. Next, the plate was centrifuged at 80*g* for 1 min to ensure cell–cell contact. After 24 h, supernatants were removed and cells were detached using accutase (Thermo Fisher Scientific), washed and stained with LIVE/DEAD Fixable Near-IR Dead Cell Stain Kit (Thermo Scientific), and the percentage of PKH26$^+$Near-IR$^+$ dead tumour cells was assessed using BD LSR Fortessa. HLA-DR expression of cancer cells was also measured using PE-Cy7 anti-human HLA-DR antibody (1:50) (BioLegend). To calculate the specific T cell-mediated cytotoxicity, target cells were separately seeded without TCC88 and their death (%) was considered as unspecific background death.

TCRVβ sequencing

To compare the TCRVβ sequence of TILs responding to the vaccine peptide pool with those responding to the bacteria-derived and microbiota-derived peptide pool, CFSE-labelled TILs were seeded at 5×10^4 cells per well in a 96-well U-bottom plates (Greiner Bio-One) and stimulated with the corresponding peptide pools incubated with irradiated autologous PBMCs (2×10^5 cells per well). After 10 days, between 5×10^3 and 10×10^3 live CD3$^+$CD4$^+$CFSE^{dim} cells were sorted from each sample. Moreover, to isolate peripheral T cells responding to bacteria-derived and microbiota-derived peptides or the IPdBP pool, CFSE-labelled cells were sorted. Sorted cells from both experiments were spun down and supernatants were removed. The pellet was snap-frozen immediately. The DNeasy Blood & Tissue Kit (Qiagen) was used to extract the DNA, and amplification and sequencing of TCRVβ CDR3 was performed using the ImmunoSEQ platform (Adaptive Biotechnologies). For the TCR data analysis, non-productive TCRVβ sequences were excluded, and shared unique productive TCRVβ sequences were specified based on identical CDR3 sequence as well as V and J chains. Representation of data was done using VENNY$^{2.1}$ (BioinfoGP, CNB-CSIC)[51].

IFNγ fluorospot assay

TILs (2×10^4 cells per well) were seeded with 5×10^4 irradiated (300 Gy) autologous EBV-transformed B cell line cells per well as APCs primed with bacteria-derived and microbiota-derived peptides (10 µM final concentration). Cells were seeded in X-VIVO 15 medium (Lonza) in human IFNγ/IL-10/IL-17A pre-coated fluorospot plates (FSP-010703-10, Mabtech) and were incubated for 44 h at 37 °C. Before seeding, fluorospot plates were washed three times using PBS and then blocked with RPMI-1640 medium (Sigma-Aldrich) containing 10% FCS (Eurobio) as instructed by the manufacturer. After 44 h of incubation, plates were washed using PBS and incubated with detection antibody, anti-IFNγ monoclonal antibody 7-B6-1-BAM, for 2 h. Next, plates were washed using PBS and incubated with anti-BAM-490 fluorophore conjugate for 1 h. Both detection antibody and fluorophore conjugate were diluted with PBS containing 0.1% BSA (Sigma-Aldrich) as instructed by the manufacturer. Next, plates were washed using PBS and incubated with fluorescence enhancer (Mabtech) for 15 min. Plates were emptied by flicking and kept in the dark for 24 h to completely dry. Finally, spot analysis was performed using the automated Mabtech IRIS fluorospot reader (Mabtech).

Immunopeptidome analyses

To analyse the immunopeptidomes of the tumour tissues and cell lines, frozen tissues and frozen cell pellets were homogenized in CHAPS-containing lysis buffer to perform standard immunoaffinity chromatography[24]. In brief, tumour tissues were washed thoroughly using PBS. Next, they were cut into small pieces and treated with digestion mix including RPMI-1640 medium (Sigma-Aldrich) containing 10% FCS (Eurobio), 1 mg ml^{-1} of collagenase IV (Sigma-Aldrich) and 0.1 mg ml^{-1} of DNase I (Sigma-Aldrich), and were incubated for 42 min at 37 °C. The reaction was stopped by adding PBS with 2 mM EDTA (Invitrogen) and the lysates were passed through a filter mesh. Next, tumour cells were isolated using Percoll (Sigma-Aldrich) gradient. HLA class II molecules were isolated using the HLA-DR-specific antibody (L243)[52] as well as the pan-HLA class II-specific antibody Tü39 (ref. [53]) (mixed 1:1, at least 1 mg or 1 mg per 1 g of tissue, both produced in-house at the Department of Immunology, Interfaculty Institute of Cell Biology, University of Tübingen) and subsequent elution of the HLA-bound peptides. Eluated peptides were

analysed by nanoflow high-performance liquid chromatography (Dionex UltimateTM 3000 Series liquid chromatography system) and tandem mass spectrometry (MS; Orbitrap Fusion Lumos, Thermo Fisher Scientific). Data-dependent acquisition of data was always conducted in technical triplicates[24]. The raw MS data files were converted into ‘mgf’ using ThermoRawFileParser[54] and analysed on MASCOT software with a non-redundant eubacteria UniProtKB/Swiss-Prot protein database (12 December 2020 with 334,492 entries) and search parameters were 0.5-Da fragment mass tolerance and 5 ppm precursor with no enzyme. All spectra were analysed manually. We only considered a spectrum valid if we clearly identified a sequence of four B-ions or Y-ions. The samples were also run on Scaffold (Scaffold version 5.2, Proteome Software) software and the ‘protein prophet’ reported false discovery rates between 2.4 and 5.4. All the acquired sequences were again re-evaluated by comparing them with UniPort human proteins, viral proteins, bacteriophage protein database[55] and genomic transposable elements (data not shown)[56]. Synthetic IPdBPs were measured with the same MS/MS method and device (Orbitrap Fusion Lumos, Thermo Fisher Scientific) and their MS/MS spectra were compared with those of the immunopeptidome peptides. As our synthetic peptides were synthesized with modifications at the N-terminal (acetylation) and C-terminal (amidation) sites, spectra often look quite different in terms of intensities and the presence or absence of B-ion and Y-ion series. The modification changed the charge of the peptides and the character of ionization patterns. The reason for synthesizing peptides with the above-end modifications stems from a convention that we previously established based on the ps-SCLs search algorithm. The latter decamer peptide mixtures, which are used for the unbiased antigen discovery studies, have N-terminal acetylation and C-terminal amidation, and consequently, we always synthesized the predicted peptides with the same ends. The peptides for the sequence validation of the IPdBPs should have been synthesized with free ends. Despite this limitation, peptide sequencing identified the exact same amino acid sequences for 37 IPdBPs via the synthetic peptides. The comparison of MS/MS spectra of ten synthetic peptides per original spectra of IPdBPs are shown in Supplementary Table 16. Universal Spectrum Explorer (USE) was used to annotate the spectra and Proteome Discoverer (Thermo Fischer) was utilized to do the research in the raw files.

The core binding motif of peptides to DRB1*03:01, DRB1*04:02, DRB3*02:02 and DRB4*01:01 was predicted using the NetMHCII 2.3 server[57], and iceLogo[58] (https://iomics.ugent.be/icelogoserver/) was used for graphical representation of the core binding motifs.

16S rRNA gene sequencing

16S rRNA gene sequencing of the tumour tissues was conducted by Microsynth AG. In brief, fresh-frozen tissues from patients were sent to Microsynth AG. Tissues were cut manually using a sterile scalpel, homogenized using Fastprep 24 (45 s, speed 6.5) and total DNA was extracted using ZymoBiomics DNA Mini Kit D4300 (Zymo Research). Negative extraction controls were included. Enrichment PCR of the whole 16S gene (primer pair 27F and 1492R) was performed, followed by amplicon library generation (Nextera library preparation protocol) using the locus-specific primer pair 341F (5′-CCT ACG GGN GGC WGC AG-3′) and 805R (5′-GAC TAC HVG GGT ATC TAA TCC-3′) and sequencing on a MiSeq platform 2 × 250 bp v2 (Illumina).

Reads were demultiplexed by separating reads into individual FastQ files, quality controlled and trimmed of Illumina adaptor sequences using locus-specific bcl2fastq software version v2.20.0.422, FastQC version 0.11.8 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) and cutadapt v3.2 (http://journal.embnet.org/index.php/embnetjournal/article/view/200), respectively. Trimmed forward and reverse reads of each paired-end read were merged to in silico reform the sequenced molecule considering a minimum overlap of 15 bases using the software USEARCH (version 11.0.667)[59]. Merged reads were denoised using the UNOISE algorithm implemented in USEARCH to form operational taxonomic units, discarding singletons and chimaeras in the process. The resulting operational taxonomic unit abundance table was then filtered for possible barcode bleed-in contaminations and against the negative controls using the UNCROSS algorithm. Operational taxonomic unit sequences were compared with the reference sequences of the RDP 16S database[60]. In addition, we used the previously reported data on the tumour microbiome[20] with in-depth contamination filtering to further remove potential bacterial contaminants.

RNA sequencing

RNA sequencing of the tumour tissue (1635WI) and analysis were conducted as previously reported[25]. In brief, RNA sequencing was conducted using the HiSeq 4000 System (Illumina) at the Functional Genomics Center Zurich. Quantile normalized gene counts were compared with The Cancer Genome Atlas (TCGA) glioblastoma cohort.

Statistical analyses

All analysis were conducted using GraphPad Prism 8.0. Data are expressed as mean ± s.e.m. Proliferation results from incorporation of ^{3}H-thymidine are always shown with SI mean ± s.e.m. and, as mentioned before, SI ≥ 2 is considered significant[26,27,32].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The TCRVB sequencing results reported in this paper are available and can be accessed via the immuneACCESS database of Adaptive Biotechnologies (https://clients.adaptivebiotech.com/pub/naghavian-2023-n). The MS immunopeptidomic raw data have been deposited to the ProteomeXchange Consortium via the PRIDE[61] partner repository with the dataset identifier PXD036811. Gene expression data for the TCGA glioblastoma cohort were downloaded from the Broad GDAC Firehose (Broad Institute of MIT and Harvard; https://doi.org/10.7908/C16T0M0N). Data were also acquired from UniProt human proteins (Swiss-Prot-20201202; https://www.uniprot.org/proteomes?query=(taxonomy_id:9606), UniProt viruses (Swiss-Prot-20210305; https://www.uniprot.org/uniprotkb?query=viruses%20AND%20(virus_host_id:9606), UniProt bacteria (Swiss-Prot-20201202; https://www.uniprot.org/proteomes/?facets=superkingdom%3ABacteria&query=%2A) and for human gut

microbiota[34]. The RNA sequencing raw data have been deposited to the European Nucleotide Archive with the accession code PRJEB56507 (ref. [25]). Source data are provided with this paper.

References

1. Routy, B. et al. Gut microbiome influences efficacy of PD-1-based immunotherapy against epithelial tumors. *Science* **359**, 91–97 (2018).

2. Vétizou, M. et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science* **350**, 1079–1084 (2015).

3. Mellman, I., Coukos, G. & Dranoff, G. Cancer immunotherapy comes of age. *Nature* **480**, 480–489 (2011).

4. Robert, C. A decade of immune-checkpoint inhibitors in cancer therapy. *Nat. Commun.* **11**, 3801 (2020).

5. Tran, E., Robbins, P. F. & Rosenberg, S. A. 'Final common pathway' of human cancer immunotherapy: targeting random somatic mutations. *Nat. Immunol.* **18**, 255–262 (2017).

6. Hodis, E. et al. A landscape of driver mutations in melanoma. *Cell* **150**, 251–263 (2012).

7. Hu, Z. et al. Personal neoantigen vaccines induce persistent memory T cell responses and epitope spreading in patients with melanoma. *Nat. Med.* **27**, 515–525 (2021).

8. Xing, Y. & Hogquist, K. A. T-cell tolerance: central and peripheral. *Cold Spring Harb. Perspect. Biol.* **4**, a006957 (2012).

9. Roberts, N. J. et al. Intratumoral injection of *Clostridium novyi*-NT spores induces antitumor responses. *Sci. Transl. Med.* **6**, 249ra111 (2014).

10. Lemos de Matos, A., Franco, L. S. & McFadden, G. Oncolytic viruses and the immune system: the dynamic duo. *Mol. Ther.* **17**, 349–358

(2020).

11. Zitvogel, L. & Kroemer, G. Cross-reactivity between cancer and microbial antigens. *Oncoimmunology* **10**, 1877416 (2021).

12. Bessell, C. A. et al. Commensal bacteria stimulate antitumor responses via T cell cross-reactivity. *JCI Insight* **5**, 135597 (2020).

13. Fluckiger, A. et al. Cross-reactivity between tumor MHC class I-restricted antigens and an enterococcal bacteriophage. *Science* **369**, 936–942 (2020).

14. Antonelli, A. C., Binyamin, A., Hohl, T. M., Glickman, M. S. & Redelman-Sidi, G. Bacterial immunotherapy for cancer induces CD4-dependent tumor-specific immunity through tumor-intrinsic interferon-γ signaling. *Proc. Natl Acad. Sci. USA* **117**, 18627–18637 (2020).

15. Kates, M. et al. Intravesical BCG Induces $CD4^+$ T-cell expansion in an immune competent model of bladder cancer. *Cancer Immunol. Res.* **5**, 594–603 (2017).

16. Linnemann, C. et al. High-throughput epitope discovery reveals frequent recognition of neo-antigens by $CD4^+$ T cells in human melanoma. *Nat. Med.* **21**, 81–85 (2015).

17. Kreiter, S. et al. Mutant MHC class II epitopes drive therapeutic immune responses to cancer. *Nature* **520**, 692–696 (2015).

18. Ott, P. A. et al. An immunogenic personal neoantigen vaccine for patients with melanoma. *Nature* **547**, 217–221 (2017).

19. Platten, M. et al. A vaccine targeting mutant IDH1 in newly diagnosed glioma. *Nature* **592**, 463–468 (2021).

20. Nejman, D. et al. The human tumor microbiome is composed of tumor type-specific intracellular bacteria. *Science* **368**, 973–980 (2020).

21. Kalaora, S. et al. Identification of bacteria-derived HLA-bound peptides in melanoma. *Nature* **592**, 138–143 (2021).

22. Riquelme, E. et al. Tumor microbiome diversity and composition influence pancreatic cancer outcomes. *Cell* **178**, 795–806.e12 (2019).

23. Galeano Niño, J. L. et al. Effect of the intratumoral microbiota on spatial and cellular heterogeneity in cancer. *Nature* **611**, 810–817 (2022).

24. Marcu, A. et al. HLA Ligand Atlas: a benign reference of HLA-presented peptides to improve T-cell-based cancer immunotherapy. *J. Immunother. Cancer* **9**, e002071 (2021).

25. Wang, J. et al. Vaccination with designed neopeptides induces intratumoral, cross-reactive $CD4^+$ T cell responses in glioblastoma. *Clin. Cancer Res.* https://doi.org/10.1158/1078-0432.CCR-22-1741 (2022).

26. Planas, R. et al. GDP-l-fucose synthase is a $CD4^+$ T cell-specific autoantigen in DRB3*02:02 patients with multiple sclerosis. *Sci. Transl. Med.* **10**, eaat4301 (2018).

27. Jelcic, I. et al. Memory B cells activate brain-homing, autoreactive $CD4^+$ T cells in multiple sclerosis. *Cell* **175**, 85–100.e23 (2018).

28. Sospedra, M. et al. Recognition of conserved amino acid motifs of common viruses and its role in autoimmunity. *PLoS Pathog.* **1**, e41 (2005).

29. Zhao, Y. et al. Combinatorial peptide libraries and biometric score matrices permit the quantitative analysis of specific and degenerate interactions between clonotypic TCR and MHC peptide ligands. *J. Immunol.* **167**, 2130–2141 (2001).

30. Pinilla, C., Appel, J. R. & Houghten, R. A. Investigation of antigen–antibody interactions using a soluble, non-support-bound synthetic decapeptide library composed of four trillion (4 × 1012) sequences. *Biochem. J.* **301**, 847–853 (1994).

31. Yousef, S. et al. TCR bias and HLA cross-restriction are strategies of human brain-infiltrating JC virus-specific $CD4^+$ T cells during viral infection. *J. Immunol.* **189**, 3618–3630 (2012).

32. Wang, J. et al. HLA-DR15 molecules jointly shape an autoreactive T cell repertoire in multiple sclerosis. *Cell* **183**, 1264–1281.e20 (2020).

33. Sospedra, M. et al. Combining positional scanning peptide libraries, HLA-DR transfectants and bioinformatics to dissect the epitope spectrum of HLA class II cross-restricted $CD4^+$ T cell clones. *J. Immunol. Methods* **353**, 93–101 (2010).

34. Almeida, A. et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat. Biotechnol.* **39**, 105–114 (2021).

35. Castellino, F. et al. Chemokines enhance immunity by guiding naive $CD8^+$ T cells to sites of $CD4^+$ T cell–dendritic cell interaction. *Nature* **440**, 890–895 (2006).

36. Wen, P. Y. et al. Glioblastoma in adults: a Society for Neuro-Oncology (SNO) and European Society of Neuro-Oncology (EANO) consensus review on current management and future directions. *Neuro Oncol.* **22**, 1073–1113 (2020).

37. Shiao, S. L. et al. Commensal bacteria and fungi differentially regulate tumor responses to radiation therapy. *Cancer Cell* **39**, 1202–1213.e6 (2021).

38. Gopalakrishnan, V. et al. Gut microbiome modulates response to anti-PD-1 immunotherapy in melanoma patients. *Science* **359**, 97–103 (2018).

39. Matson, V. et al. The commensal microbiome is associated with anti-PD-1 efficacy in metastatic melanoma patients. *Science* **359**, 104–108 (2018).

40. Fu, A. et al. Tumor-resident intracellular microbiota promotes metastatic colonization in breast cancer. *Cell* **185**, 1356–1372.e26

(2022).

41. Fritsche, J. et al. Pitfalls in HLA ligandomics—how to catch a li(e)gand. *Mol. Cell. Proteomics* **20**, 100110 (2021).

42. Nakanishi, Y., Lu, B., Gerard, C. & Iwasaki, A. $CD8^+$ T lymphocyte mobilization to virus-infected tissue requires $CD4^+$ T-cell help. *Nature* **462**, 510–513 (2009).

43. Ayers, M. et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J. Clin. Invest.* **127**, 2930–2940 (2017).

44. Vujanovic, L., Mandic, M., Olson, W. C., Kirkwood, J. M. & Storkus, W. J. A mycoplasma peptide elicits heteroclitic $CD4^+$ T cell responses against tumor antigen MAGE-A6. *Clin. Cancer Res.* **13**, 6796–6806 (2007).

45. Chiou, S.-H. et al. Global analysis of shared T cell specificities in human non-small cell lung cancer enables HLA inference and antigen discovery. *Immunity* **54**, 586–602.e8 (2021).

46. Lee, C. H. et al. Predicting cross-reactivity and antigen specificity of T cell receptors. *Front. Immunol.* **11**, 565096 (2020).

47. Wooldridge, L. et al. A single autoimmune T cell receptor recognizes more than a million different peptides. *J. Biol. Chem.* **287**, 1168–1177 (2012).

48. Galperin, M. et al. $CD4^+$ T cell-mediated HLA class II cross-restriction in HIV controllers. *Sci. Immunol.* **3**, eaat0687 (2018).

49. Zuccarini, M. et al. The role of Wnt signal in glioblastoma development and progression: a possible new pharmacological target for the therapy of this tumor. *Genes (Basel)* **9**, 105 (2018).

50. Hosang, L. et al. The lung microbiome regulates brain autoimmunity. *Nature* **603**, 138–144 (2022).

51. Oliveros, J. C. Venny: an interactive tool for comparing lists with Venn's diagrams. *BioinfoGP* https://bioinfogp.cnb.csic.es/tools/venny/ (2007).

52. Shackelford, D. A., Lampson, L. A. & Strominger, J. L. Separation of three class II antigens from a homozygous human B cell line. *J. Immunol.* **130**, 289–296 (1983).

53. Odum, N. et al. Homotypic aggregation of human cell lines by HLA class II-, class Ia- and HLA-G-specific monoclonal antibodies. *Eur. J. Immunol.* **21**, 2121–2131 (1991).

54. Hulstaert, N. et al. ThermoRawFileParser: modular, scalable, and cross-platform RAW file conversion. *J. Proteome Res.* **19**, 537–542 (2020).

55. Camarillo-Guerrero, L. F., Almeida, A., Rangel-Pineros, G., Finn, R. D. & Lawley, T. D. Massive expansion of human gut bacteriophage diversity. *Cell* **184**, 1098–1109.e9 (2021).

56. Bonté, P.-E. et al. Single-cell RNA-seq-based proteogenomics identifies glioblastoma-specific transposable elements encoding HLA-I-presented peptides. *Cell Rep.* **39**, 110916 (2022).

57. Jensen, K. K. et al. Improved methods for predicting peptide binding affinity to MHC class II molecules. *Immunology* **154**, 394–406 (2018).

58. Colaert, N., Helsens, K., Martens, L., Vandekerckhove, J. & Gevaert, K. Improved visualization of protein consensus sequences by iceLogo. *Nat. Methods* **6**, 786–787 (2009).

59. Edgar, R. C. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**, 2460–2461 (2010).

60. Cole, J. R. et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res.* **42**, D633–D642 (2014).

61. Perez-Riverol, Y. et al. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* **47**, D442–D450 (2019).

62. Rammensee, H., Bachmann, J., Emmerich, N. P., Bachor, O. A. & Stevanović, S. SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics* **50**, 213–219 (1999).

Acknowledgements

Study support came from a European Research Council grant (ERC-2013-ADG-340733) and a Swiss National Science Foundation grant (SNF 320030_153213). We thank I. Jelcic, T. Lorenzini, M. Wawrzyniak and L. Grob for their help; A. Lutterotti, R. Planas, P. Tomas Ojer, T. Weiss, P. Roth, H. Moch, E. Rushing, S. Dettwiler, B. Schrörs, J. H. Shin, R. McKay, C. J. Wu and E. F. Greiner who participated in the personalized vaccination[25]; and C. Falkenburger, U. Wulle, P. Hrstić, N. Bauer and B. Pömmerl for excellent technical support. Figures 1a and 2a and Extended Data Fig. 3b were created with BioRender (https://biorender.com).

Funding

Open access funding provided by University of Zurich.

Author information

Author notes

1. These authors contributed equally: Mireia Sospedra, Marian C. Neidert

Authors and Affiliations

1. Neuroimmunology and MS Research Section (NIMS), Neurology Clinic, University of Zurich, University Hospital Zurich, Zurich, Switzerland

   Reza Naghavian, Wolfgang Faigle, Pietro Oldrati, Jian Wang, Yuhan Qiu, Mireia Sospedra & Roland Martin

2. Cellerys AG, Schlieren, Switzerland

Reza Naghavian, Wolfgang Faigle, Mireia Sospedra & Roland Martin

3. Immunity and Cancer, Institut Curie, PSL University, INSERM U932, Paris, France

Wolfgang Faigle, Pierre-Emmanuel Bonté & Sebastian Amigorena

4. School of Basic Medical Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, China

Jian Wang

5. NEXUS Personalized Health Technologies, ETH Zurich, Schlieren, Switzerland

Nora C. Toussaint

6. Swiss Institute of Bioinformatics, Zurich, Switzerland

Nora C. Toussaint

7. Clinical Neuroscience Center, Department of Neurosurgery, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Gioele Medici, Luca Regli & Marian C. Neidert

8. Department of Peptide-based Immunotherapy, University of Tübingen, University Hospital Tübingen, Tübingen, Germany

Marcel Wacker & Juliane S. Walz

9. Institute for Cell Biology, Department of Immunology, University of Tübingen, Tübingen, Germany

Marcel Wacker, Lena K. Freudenmann, Hans-Georg Rammensee & Juliane S. Walz

10. German Cancer Consortium (DKTK) and German Cancer Research Center (DKFZ), partner site Tübingen, Tübingen, Germany

Marcel Wacker, Lena K. Freudenmann, Hans-Georg Rammensee & Juliane S. Walz

11. Cluster of Excellence iFIT (EXC 2180) 'Image-Guided and Functionally Instructed Tumor Therapies', University of Tübingen, Tübingen, Germany

Marcel Wacker, Lena K. Freudenmann, Hans-Georg Rammensee & Juliane S. Walz

12. Laboratory of Molecular Neuro-Oncology, Department of Neurology and Clinical Neuroscience, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Michael Weller

13. Clinical Collaboration Unit Translational Immunology, German Cancer Consortium (DKTK), Department of Internal Medicine, University Hospital Tübingen, Tübingen, Germany

Juliane S. Walz

14. Department of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland

Silvio D. Brugger

15. Department of Neurosurgery, University Hospital Hamburg Eppendorf, University of Hamburg, Hamburg, Germany

Malte Mohme

16. Computational and Systems Biology Branch, Biometric Research Program, Division of Cancer Treatment and Diagnosis, NCI, NIH, Rockville, MD, USA

Yingdong Zhao

17. Department of Neurosurgery, Cantonal Hospital St. Gallen, St. Gallen, Switzerland

    Marian C. Neidert

18. Institute of Experimental Immunology, University of Zurich, Zurich, Switzerland

    Roland Martin

19. Therapeutic Immune Design Unit, Center for Molecular Medicine, Department of Clinical Neurosciences, Karolinska Institutet, Stockholm, Sweden

    Roland Martin

Contributions

R.N. designed, performed and analysed experiments (proliferation assays, cell sorting, cytotoxicity testing, TCR sequencing analysis, T cell cloning, cytokine and ELISA analysis) and contributed to writing the manuscript. W.F. performed immunopeptidome annotations with bacteria and microbial databases. J.W. participated in experimental design and performed T cell cloning. P.O., N.C.T. and Y.Z. provided protein databases, the tool for peptide searching and performed target searches. M.C.N. provided tumour tissue and cancer cells for immunopeptidome analyses, and participated in designing the immunopeptidome and cytotoxicity experiments and discussing the manuscript. L.K.F. and H.-G.R. performed immunopeptidome analyses and commented on the manuscript. M.Wacker and J.S.W. provided the MS/MS fragmentation of synthetic peptides and the analyses. G.M. provided patient-derived cancer cells for the cytotoxicity experiments and participated in designing the assay. S.D.B. assisted with bacterial taxonomy and 16S rRNA sequencing. M.Weller provided tumour tissue and patient-derived cancer cells and immunopeptidome data, participated in designing patient cohort and logistics, and reviewed the manuscript. L.R. provided tumour tissue, participated in designing the patient cohort and logistics, and

reviewed the manuscript. M.M. provided additional glioblastoma-derived TILs and tumour tissues, and participated in designing the experiment. P.-E.B. and S.A. contributed with annotation of the immunopeptidome with non-coding sequences. Y.Q. performed expansion of cells and testing with autoantigen peptide pools. M.S. provided expertise in epitope mapping and interpreting the scoring matrix data. R.M. supervised and designed the overall study, interpreted the data and wrote the manuscript.

Corresponding author

Correspondence to Roland Martin.

Ethics declarations

Competing interests

R.N., W.F. and M.S. are currently employees of Cellerys AG. R.M. received unrestricted grants from Biogen, Novartis, Roche, Third Rock; has advisory roles and lectures for Roche, Novartis, Biogen, Merck, Genzyme, Neuway, CellProtect, Third Rock and Teva; is listed as an inventor on a NIH-held patent of daclizumab in MS and University of Zurich-held patents on JCV VP1 for vaccination against PML, JCV-specific neutralizing antibodies to treat PML, antigen-specific tolerization with peptide-coupled cells and novel autoantigens in MS, the use of designer neoantigens for glioblastoma (inventor R.M.), and of peptides and TCRs shown in this study for the personalized treatment of glioblastoma (inventors R.N., J.W. & R.M.); and is a co-founder of Abata, and a co-founder and employee of Cellerys AG. L.K.F. is an employee of Immatics Biotechnologies. M.C.N. has received a research grant from Novocure, and honoraria for consulting or lectures from WISE, MSD and Novocure. M.Weller has received research grants from Quercis and Versameb, and honoraria for lectures or advisory board participation or consulting from Bayer, Curevac, Medac, Merck (EMD), Novartis, Novocure, Orbus, Philogen, Roche and Sandoz.

Peer review

Peer review information

Nature thanks Laurence Zitvogel and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Confirming the presence of bacteria in tumor tissues (see also Fig. 1, Extended Data Fig. 2 and Supplementary Table 4).

a, HLA-II immunopeptidome analyses for bacterial peptides in control including brain lesions of 3 MS patients and brain tissue of 6 healthy donors. A total of 55 bacterial peptides was identified from all the samples. Identical bacterial peptides in the control cohort were removed from the GBM tumor-specific bacterial peptides. **b**, 16S rRNA gene sequencing confirmed bacteria in tumor tissues of 10 GBM patients, from whom immunopeptidome analyses were available. OTUs relative abundance in read counts were shown at the phylum level for each tumor (Supplementary Tables 5 and 6). * 16S rRNA data from patient 1635WI includes both primary and recurrent tumors of the patient in triplicates. OTUs relative abundance was calculated based on mean OTU of the triplicates divided by total OTUs in each tumor. **c**, Deduced core binding motif all HLA-DR-derived peptides annotated to human/tumor proteins and UniProt bacterial peptides from primary and recurrent tumors (patient 1635WI) were compared using iceLogo ($p = 0.05$, two-tailed independent *t*-test by IceLogo[58]). Light blue squares show the known anchor positions for corresponding HLA-DR molecules[62].

Extended Data Fig. 2 HLA-II expression of tumor tissue and cancer cell line.

a, RNA-seq analysis of the tumor tissue (1635WI) was conducted by TCGA-RNAseqv2 pipeline (https://webshare.bioinf.unc.edu/public/mRNAseq_TCGA/UNC_mRNAseq_summary.pdf)[25]. Normalized gene counts of HLA-DR were compared with The Cancer Genome Atlas (TCGA) GBM cohort. White square represents normal expression of the gene in the tumor (Supplementary Table 14). **b**, Cancer cells isolated from the tumor tissue (patient 1635WI) and expanded *in vitro* were stained with HLA-DR antibody. **c**, Paraffin-embedded recurrent tumor tissue (patient 1635WI) was stained for GFAP,

CD31, CD3, CD68 and HLA-DR[25]. HLA-DR is present on the surface of immune cells (CD3+), endothelial cells (CD31) and GFAP+ astrocytes (6 different regions in n = 1 section of the tumor were used for this analysis). Scale bar: 200 μm.

Extended Data Fig. 3 T cell cloning using SIN3A* tumor antigen and graphical summary of TCC isolation, response of TCC88 to ps-SCL, and depiction of database screening for antigen discovery approach (see also Figs. 1 and 2 and Extended Data Fig. 4).

a, Gating strategy for isolation of TCCs from TILs responding to SIN3A* peptide. 60 replicate wells of CFSE-labeled TILs were pooled after 10 days stimulation with SIN3A* tumor antigen. Cells were stained using live-dead, CD3, CD4 and were next single-cell sorted. **b**, Schematic depiction of the steps for discovering target antigens for a TCC. TCCs are tested with positional scanning combinatorial peptide libraries (200 mixtures of decamer peptides) presented by BLS cells expressing a single HLA-DR molecule, which had been previously established as the restriction element for each TCC. Proliferative or cytokine responses of TCC to each of the decamer mixtures, i.e. 200 compounds are measured. Data are then used to create a scoring matrix. Antigens with predicted high stimulatory potential are identified by searches of several databases including human proteome, autologous tumor transcriptome and peptidome, pathogenic bacteria and viruses as well as human gut microbiota. Candidate peptides originating from pathogenic bacteria and gut microbiota strongly stimulated TCC88. Further, these peptides were used to stimulate other tumor-reactive TCCs as well as bulk TILs and peripheral memory $CD4^{+}$ T cells. The schematics in panel **b** were created using BioRender (https://biorender.com).

Extended Data Fig. 4 (see also Fig. 3). HLA-DR restriction and response to autoantigens of TCC88.

a, Several of the isolated TCCs were stimulated with irradiated BLS cells expressing single HLA-DR molecules of patient 1635WI (BLS-DRB1*03:01, BLS-DRB1*04:02, BLS-DRB3*02:02 and BLS-

DRB4*01:01) as APCs and dose titration of SIN3A* peptide (10, 1, 0.1, 0.01 and 0.001 μM). Proliferation was measured using ^{3}H-thymidine incorporation after 3 days. Data represents mean SI of 5 replicate wells ± SEM). Several TCCs did not respond to peptides presented on any of HLA-DR molecules (data not shown). **b**-**c**, Median IFN-γ response of three independent experiments (transformed to log base 10 values) of TCC88 stimulated via ps-SCL was measured and depicted for 20 L-amino acids fixed in all positions of a 10-mer peptide. **c**, IFN-γ values were depicted in a table format (scoring matrix) and represent the relative stimulatory potency of each L-amino acid in different positions of the decamers. **d**, Relationship between predicted score and stimulatory index (SI) was established by testing potential cognate antigens within the predicted score range of 60–100% of the maximal score with TCC88. **e**–**f**, TCC88 was tested with pools of overlapping peptides (including 260 peptides) covering brain, pancreas and skin autoantigens (Supplementary Table 11) including peptide pools covering myelin basic protein (MBP pools), myelin oligodendrocyte glycoprotein (MOG pools), myelin proteolipid protein (PLP pool contains 5 immunodominant PLP peptides), RAS guanyl-releasing protein 2 (RASGRP2 pools),GDP-L-fucose synthase (TSTA3 pools), MELAN-A, tyrosinase (TYR), premelanosome protein (PMEL), insulin, SLC30A8 (ZnT8) and islet amyloid polypeptide (IAPP). TCC88 was stimulated with peptide-pulsed irradiated BLS-DRB3*02:02 cells for 3 days. Proliferation was measured using ^{3}H-thymidine incorporation (data represents mean value of 3 replicate wells ±SEM). **f**, Response of TCC88 to unmutated SIN3A peptide was compared to SIN3A* at different concentrations including 10, 1, 0.1, 0.01 and 0.001 μM.

Source data

Extended Data Fig. 5 Response of TILs and TCCs to bacteria/microbiota-derived peptides (see also Figs. 3 and 4).

a, The source bacterium of bacteria/microbiota-derived peptides that stimulate TCC88 is shown at the phylum and class levels. Source proteins of the peptides were used to check within UniProt databases and taxonomy data was then extracted. **b**, Heatmap result analyzing all peptides that were stimulatory or non-stimulatory for TCC88 with the aim to examine if amino

acid motifs are present in stimulatory peptides (Green, black and red represent higher, average and lower prevalence of distinct amino acids in the positions of the tested 10-mer peptides. IceLogo[58] utilizes the z-score to calculate amino acid specific p-values at each position via an error function[58]). **c**, Supernatants were collected from bulk TILs stimulated with bacteria/microbiota-derived peptides after 5 days. IFN-γ secretion of TILs was measured by ELISA (data represents mean value of 5 replicate wells ± SEM). **d**, Several TIL-derived TCCs were stimulated with bacteria/microbiota-derived peptides and irradiated autologous PBMCs as APCs. Proliferation was measured by ^{3}H-thymidine incorporation after 3 days (data represents mean value of 2 replicate wells). **e**, Supernatants were collected after stimulation of TCC75 with SIN3A* and HGM27 peptides for 3 days. Cytokine/chemokine secretion of TCC75 was examined for the different antigens and depicted in a heatmap. Specific amount of each cytokine/chemokine (pg/ml) was calculated by subtracting the cytokines in the negative control cultures (no peptide control).

Source data

Extended Data Fig. 6 TILs depleted of TCC88 and also other TCCs respond to bacteria/microbiota-derived peptides (see also Figs. 3 and 4).

a, Gating strategy for bulk TILs using CD4, CD8, PE- and FITC-TCRVβ antibodies. Two populations of CD4$^+$ PE-TCRVβ$^+$ and CD4$^+$ FITC-TCRVβ$^+$ were sorted and later mixed together. Unstained TILs were used as control to demonstrate specific staining of TILs with TCRVβ antibodies of interest. **b**, Sorted CD4$^+$ PE and FITC TCRVβ$^+$ were mixed and rested. Cells were then stimulated with irradiated autologous PBMCs pulsed with several bacteria/microbiota-derived peptides at 10 μM and proliferation was measured after 3 days using ^{3}H-thymidine incorporation (data represents mean value of 3 replicate wells ± SEM). **c**, Supernatants were tested for secretion of a broader range of cytokines/chemokines. Cytokine/chemokine secretion of bacteria/microbiota-reactive TILs was measured after stimulation with vaccine and bacteria/microbiota-derived peptide pools (left panel). Cytokine/chemokine secretion of bacteria/microbiota-reactive

peripheral $CD4^+$ T cells (selected peripheral $CD4^+$) and $CD45RA^-$ peripheral T cells (unselected peripheral $CD4^+$) in response to a tumor antigen (LRP6) was measured (right panel). Specific amount of each cytokine/chemokine (pg/ml) was calculated by subtracting the cytokines in the negative control cultures (no peptide control).

Source data

Extended Data Fig. 7 Stimulatory effect of IPdBPs on peripheral memory T cells (see also Fig. 5).

a, $CD45RA^-$ PBMCs from patient 1635WI were CFSE-labeled and stimulated with a pool of IPdBPs (see Supplementary Table 13). 60 replicate wells were pooled after 10 days and stained with CD3, CD4 and CD8. Proliferating ($CFSE^{dim}$) $CD4^+$ T cells were sorted and short-term expanded. **b**, Expanded and IPdBP-pre-stimulated peripheral $CD4^+$ T cells were tested with 10 µM of autologous tumor peptides derived from over-expressed or mutated proteins (Supplementary Table 13). Proliferation was measured using ^{3}H-thymidine incorporation (data represents mean value of 4 replicate wells ±SEM). GM-CSF ELISA was used as another read-out for T cell activation. GM-CSF was measured by ELISA and is depicted in (pg/ml) after subtraction of the negative control. The peptide pool that was used to pre-stimulate peripheral blood CD4+ T cells (IPdBP pool) served as a positive control (see material and methods section). **c**, Cytokine/chemokine response of peripheral memory $CD4^+$ T cells that had been pre-stimulated with the pool of IPdBPs and short-term expanded *in vitro*, to the autologous tumor peptide (LRP6) as well as the IPdBP pool was measured. Specific amount of each cytokine/chemokine (pg/ml) was calculated by subtracting the cytokines in the negative control cultures (no peptide control). **d**–**e**, Percentage and frequency overlap of unique productive TCRVβ sequences of $CFSE^{dim}$ peripheral memory $CD4^+$ T cells responding to IPdBP pool with TCRVβ sequences of bulk TILs. Overlap of unique productive TCRVβ sequences of IPdBP-reactive peripheral $CD4^+$ ($CFSE^{dim}$) (**d**) and bacteria/microbiota-reactive peripheral $CD4^+$ ($CFSE^{dim}$) (**e**).

Source data

Extended Data Fig. 8 IFN-γ fluorospot testing of expanded TILs from patient GBM-40.

a, 64 bacteria/microbiota-derived peptides that were stimulatory for TILs from patient 1635WI were tested with TILs from GBM-40 patient. TILs were seeded with irradiated autologous BLCLs primed with bacteria/microbiota-derived peptides and in triplicates. ΔSFU (spot forming units) for each peptide was calculated after subtraction of the mean number of spots in negative control (data represents mean value of 3 replicate wells ± SEM). **b**, Positive and negative responses of TILs to several bacteria/microbiota-derived peptides are shown on the lower panel. CD2-CD3-CD28 beads are used as positive control.

Source data

Extended Data Fig. 9 Summary of T cell specificities, TCR sharing and comparison of bacterial species identified via immunopeptidome analyses and ps-SCL screening of TCC88.

a, Overview of proliferative responses of all tested TCCs and bulk T cell populations to different antigens. Stimulatory antigens are depicted in red, non-stimulatory antigens in blue (N/A means that T cells were not tested with the respective antigens, * Peripheral memory $CD4^+$ T cells selected by IPdBPs, ** Peripheral memory $CD4^+$ T cells selected by bacteria/microbiota-derived peptides). **b**. Overview of shared TCRs between T cells from different compartments. Green displays the presence of shared TCRs (N/A means that samples were not examined for presence of shared TCRs, * Peripheral memory $CD4^+$ T cells selected by IPdBPs, ** Peripheral memory $CD4^+$ T cells selected by bacteria/microbiota-derived peptides). **c**. Comparison of peptides from bacterial species identified via HLA-II immunopeptidome analysis versus those derived from unbiased ps-SCL screening of TCC88. Bars of blue and red for *Bacillus sp*. and *Bacteroides ovatus* mean that these species were found via both approaches.

Supplementary information

Supplementary Tables

This file contains Supplementary Tables 1, 2, 9, 12, 13 and 16.

Reporting Summary

Peer Review File

Supplementary Table 3

Supplementary Table 4

Supplementary Table 5

Supplementary Table 6

Supplementary Table 7

Supplementary Table 8

Supplementary Table 10

Supplementary Table 11

Supplementary Table 14

Supplementary Table 15

Source data

Source Data Fig. 1

Source Data Fig. 2

Source Data Fig. 3

Source Data Fig. 4

Source Data Fig. 5

Source Data Extended Data Fig. 4

Source Data Extended Data Fig. 5

Source Data Extended Data Fig. 6

Source Data Extended Data Fig. 7

Source Data Extended Data Fig. 8

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Naghavian, R., Faigle, W., Oldrati, P. *et al.* Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma. *Nature* **617**, 807–817 (2023). https://doi.org/10.1038/s41586-023-06081-w

- Received: 07 June 2022
- Accepted: 13 April 2023
- Published: 17 May 2023
- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-06081-w

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

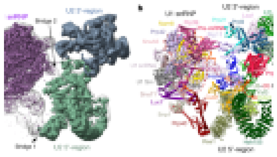
Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

---

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06081-w

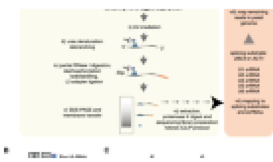
---



Structural insights into how Prp5 proofreads the pre-mRNA branch site

04 August 2021

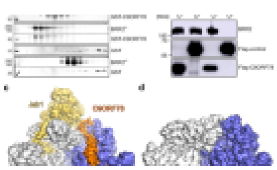
Zhenwei Zhang, Norbert Rigo, … Reinhard Lührmann



psiCLIP reveals dynamic RNA binding by DEAH-box helicases before and after exon ligation

05 March 2021

Lisa M. Strittmatter, Charlotte Capitanchik, … Kiyoshi Nagai



A multi-factor trafficking site on the spliceosome remodeling enzyme BRR2 recruits C9ORF78 to regulate alternative splicing

03 March 2022

Alexandra Bergfort, Marco Preußner, … Markus C. Wahl



Prespliceosome structure provides insights into spliceosome assembly and regulation

11 July 2018

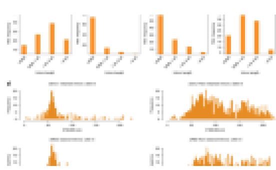
Clemens Plaschka, Pei-Chun Lin, … Kiyoshi Nagai



Structures of the human pre-catalytic spliceosome and its precursor spliceosome

12 October 2018

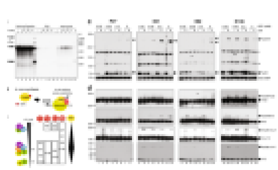
Xiechao Zhan, Chuangye Yan, … Yigong Shi



Smu1 and RED are required for activation of spliceosomal B complexes assembled on short introns

13 August 2019

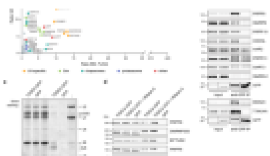
Sandra Keiper, Panagiotis Papasaikas, … Reinhard Lührmann



Activation of Prp28 ATPase by phosphorylated Npl3 at a critical step of spliceosome remodeling

25 May 2021

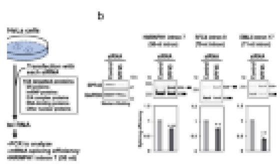
Fu-Lung Yeh, Shang-Lin Chang, … Tien-Hsien Chang



TSSC4 is a component of U5 snRNP that promotes tri-snRNP formation

15 June 2021

Klára Klimešová, Jitka Vojáčková, … David Staněk



SPF45/RBM17-dependent, but not U2AF-dependent, splicing in a distinct subset of human short introns

13 August 2021

Kazuhiro Fukumura, Rei Yoshimoto, … Akila Mayeda

- Article
- Open Access
- Published: 10 May 2023

Structural basis of catalytic activation in human splicing

- Jana Schmitzová,
- Constantin Cretu,
- Christian Dienemann,
- Henning Urlaub &
- …
- Vladimir Pena

Show authors

Nature **volume 617**, pages 842–850 (2023)

- 6917 Accesses
- 1 Citations
- 41 Altmetric
- Metrics details

Abstract

Pre-mRNA splicing follows a pathway driven by ATP-dependent RNA helicases. A crucial event of the splicing pathway is the catalytic activation, which takes place at the transition between the activated B^{act} and the branching-competent B^* spliceosomes. Catalytic activation occurs through an ATP-dependent remodelling mediated by the helicase PRP2 (also known as DHX16)[1,2,3]. However, because PRP2 is observed only at the periphery of spliceosomes[3,4,5], its function has remained elusive. Here we show that catalytic activation occurs in two ATP-dependent stages driven by two helicases: PRP2 and Aquarius. The role of Aquarius in splicing has been enigmatic[6,7]. Here the inactivation of Aquarius leads to the stalling of a spliceosome intermediate—the B^{AQR} complex—found halfway through the

catalytic activation process. The cryogenic electron microscopy structure of B^{AQR} reveals how PRP2 and Aquarius remodel B^{act} and B^{AQR}, respectively. Notably, PRP2 translocates along the intron while it strips away the RES complex, opens the SF3B1 clamp and unfastens the branch helix. Translocation terminates six nucleotides downstream of the branch site through an assembly of PPIL4, SKIP and the amino-terminal domain of PRP2. Finally, Aquarius enables the dissociation of PRP2, plus the SF3A and SF3B complexes, which promotes the relocation of the branch duplex for catalysis. This work elucidates catalytic activation in human splicing, reveals how a DEAH helicase operates and provides a paradigm for how helicases can coordinate their activities.

Main

Pre-mRNA splicing occurs in two transesterification reactions on the spliceosome, the sequential assembly and remodelling of which are driven by eight conserved ATP-dependent RNA helicases[4,8,9,10]. The branching reaction occurs between the branch site adenosine (BS-A) and the 5′ splice site (5′SS) during the formation of the post-branching complex (also known as the C complex). As a prerequisite for branching, the reactants juxtapose to the catalytic centre in a process known as catalytic activation, which occurs at the conversion of the B^{act} (also known as activated spliceosome) to the branching-competent B* (also known as catalytically activated spliceosome) complex[1]. The B∗ and the post-branching C complexes are very similar in structure and composition, differing primarily in the phosphodiester bond between the BS-A and the 5'SS[10].

Catalytic activation requires the remodelling of B^{act} spliceosomes. As part of the U2–BS duplex (also known as the branch duplex), the BS-A is sequestered 50 Å away from the catalytic centre and is enclosed by the SF3B complex of the U2 small nuclear ribonucleoprotein (snRNP)[2,3]. This inactive conformation of B^{act} is stabilized by the RES complex and the proteins SF3B1 and PRP8. Additionally, the proteins SF3A2 and CWC24 (also known as RNF113A) shield the 5′SS from the catalytic centre. These contacts must be disrupted during catalytic activation to relocate the branch helix to the active centre for catalysis[2,11,12,13].

In *Saccharomyces cerevisiae* (budding yeast), the DEAH helicase Prp2p drives catalytic activation by releasing RES, SF3a and SF3b complexes, and the proteins Cwc24p and Cwc27p; however, mechanistic details of this process remain largely unknown[1,5,14]. In humans, in addition to PRP2 (an orthologue of Prp2p), the helicase Aquarius is recruited to the spliceosome as part of a pentameric intron-binding complex (IBC), which contributes to the B-to-C transitions of the spliceosome. The spliceosome remodelling promoted by Aquarius and the time point of its action relative to PRP2 is unknown[6,7]. Aquarius also stands out as the only splicing helicase from the SF1 family.

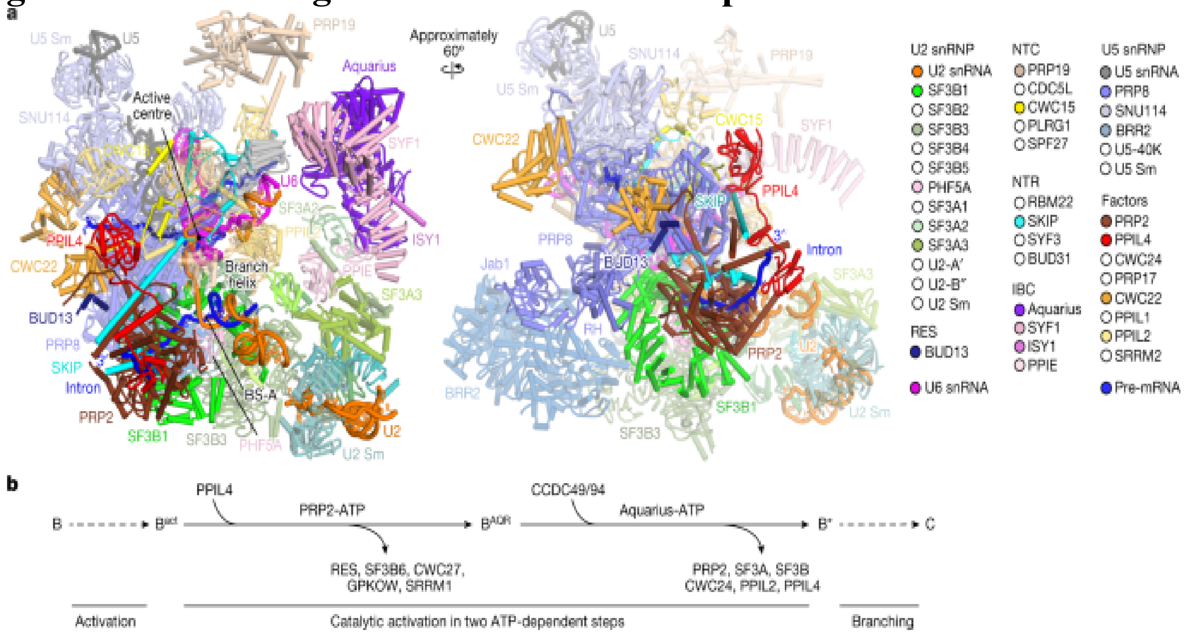
Here we reconstitute a spliceosome stalled between the actions of PRP2 and Aquarius, which we refer to as the B^{AQR} complex. The cryogenic electron microscopy (cryo-EM) structure of the B^{AQR} spliceosome reveals the mechanism of B^{act}-to-B remodelling and the mode of action of PRP2 assisted by interacting proteins. Our results also provide insight into how the two helicases coordinate their activities to promote the branching reaction.

Overview of the human B^{AQR} spliceosome

After identifying the IBC as a complex that delivers Aquarius to the spliceosome[6], we reconstituted the B^{AQR} spliceosome on MINX pre-mRNA in HeLa nuclear extracts supplemented with IBC carrying the dominant-negative Aquarius mutant K829A (IBC(K829A)). This mutant stalls the spliceosome before branching[6] (Extended Data Fig. 1a–d and Supplementary Fig. 1). The proteome of B^{AQR} resembled one of the B^{act} complexes[2,13,15]. However, within B^{AQR}, PRP2 and PPIL4 were abundant, whereas the PRP2 cofactor GPKOW was detected in low amounts (Supplementary Data 1).

We reconstructed a cryo-EM map of the B^{AQR} core complex at about 2.9 Å resolution. Focused 3D classification allowed us to resolve the peripheral regions of the complex and model in the BRR2 helicase, the PRP19 and IBC complexes, and the 3′ module of the U2 snRNP (Fig. 1a, Extended Data Figs. 1e–i, 2 and 3a–d, Extended Data Table 1, Supplementary Data 2 and Supplementary Video 1).

Fig. 1: Structural organization of the B^{AQR} spliceosome.



a, Overview of the B^{AQR} cryo-EM structure. Key subunits are colour coded. **b**, Compositional remodelling of the human spliceosome during catalytic activation. Subunits recruited or destabilized by the ATPase activities of PRP2 and Aquarius are indicated. Destabilized subunits can remain flexibly attached to the spliceosomes, often at lower stoichiometry. NTR, NTC-related proteins.

Compared with B^{act} complexes[2,13], B^{AQR} exhibited repositioning of PRP2, the SF3B complex, SKIP (also known as PRP45), the Jab1 domain and the RNaseH-like (RH) domain of PRP8, and the intron RNA downstream of the branch helix. More subtle changes were present for BRR2, CWC15 (also known as AD002) and CWC22. Furthermore, PPIL4 was recruited to the B^{AQR} complex while the RES complex, SF3B6 (also known as p14), CWC27, SRRM1 and GPKOW were destabilized (Fig. 1 and Supplementary Video 1).

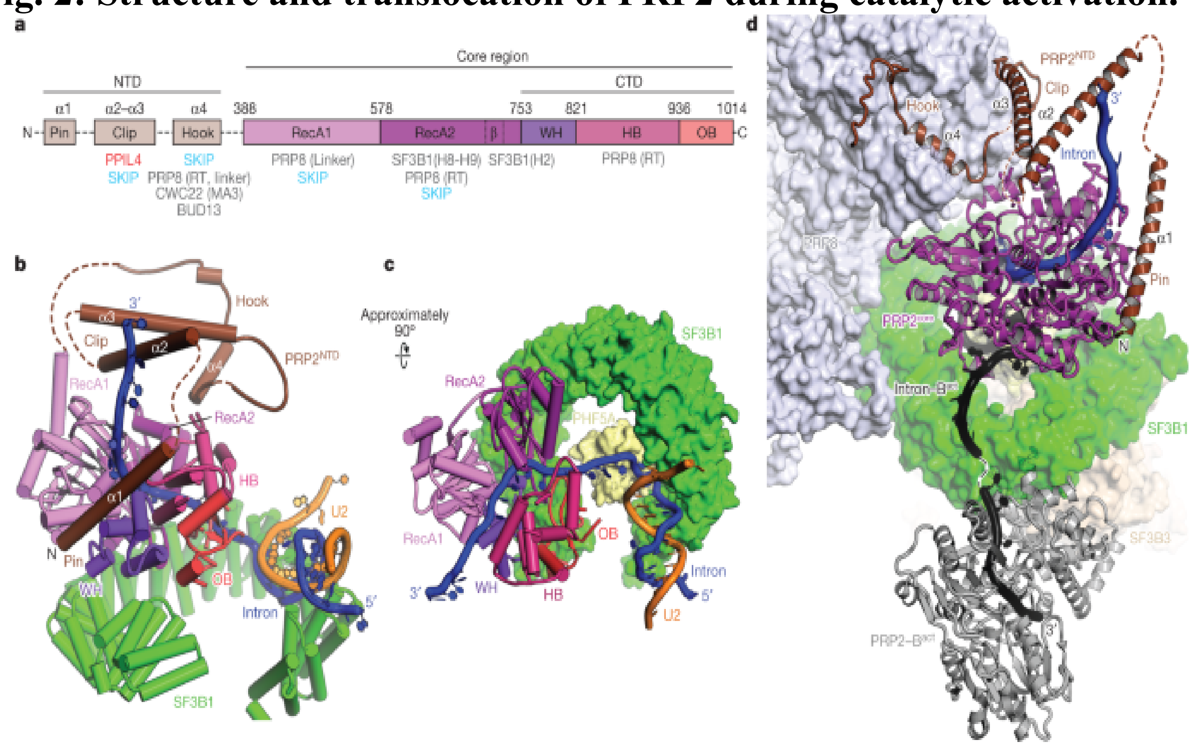
The position of Aquarius was primarily the same in B^{act} complexes and B^{AQR} complexes[2,13], but notably different in C complexes[2,16], which indicates that inactivated Aquarius prevents the conversion of B^{AQR}-to-C complexes (Extended Data Fig. 3e). By contrast, PRP2 occupied different locations in B^{act} and B^{AQR}. This result indicates that B^{AQR} is an intermediate

spliceosome stalled after the translocation of PRP2 but before the action of Aquarius. We conclude that B^{AQR} represents an intermediate of the splicing pathway that follows the B^{act} complex while preceding B^{*} and C complexes[1,11,12,17] (Fig. 1b).

Structure of PRP2 in the B^{AQR} spliceosome

PRP2 is composed of a flexible N-terminal region (residues 1–387) and a conserved globular core ($PRP2^{core}$, residues 388–1042), which is typical of all four spliceosomal DEAH helicases (Fig. 2a). Structures of fungal $PRP2^{core}$ have been reported for *Chaetomium thermophilum* and budding yeast in complex with ligands such as ADP, ADP-BeF_3^-, RNA oligonucleotides and the cofactor Spp2p (refs. [5,18]). The $PRP2^{core}$ has also been resolved at 3.2 Å resolution in yeast B^{act} spliceosomes or docked as a homology model in human B^{act} complexes[5,11,13].

Fig. 2: Structure and translocation of PRP2 during catalytic activation.



a, Domain composition and interactions of PRP2 in the B^{AQR} complex. **b**,**c**, Structure and conformation of PRP2, the intron and SF3B1 in B^{AQR}. Note

that PRP2^{NTD} (the N-terminal domain) is not visible in **c**. **d**, In B^{AQR}, PRP2 is moved in a cavity framed by PRP8 and SF3B1. The previous location and conformation of PRP2 and the intron in B^{act} are shown in grey and black, respectively.

The cryo-EM density map of human B^{AQR} enabled the building of the PRP2^{core} and of the accessory N-terminal domain (PRP2^{NTD}, residues 265–295; Fig. 2b,c and Extended Data Fig. 3c,d). The PRP2^{core} region exhibited the two RecA-like domains (RecA1 and RecA2) and a carboxy-terminal module (PRP2^{CTD}) that encompassed the winged-helix (WH), oligonucleotide/oligosaccharide-binding (OB) and helix-bundle (HB) domains. The PRP2^{core} from B^{AQR} is in the open, post-ATP hydrolytic conformation. This conformation has previously only been observed for the ATPase-defective mutant Prp2p(K252A) in the yeast B^{act} spliceosome[5].

The structure and function of PRP2^{NTD} were unknown. In the B^{AQR} spliceosome, PRP2^{NTD} acquired a stage-defining fold that lacked a hydrophobic core and was organized into three distinct modules. Owing to their extended conformation, they interacted with several spliceosomal subunits. We refer to these elements as the hook, the clip and the pin (Fig. 2a,b and Extended Data Fig. 4a).

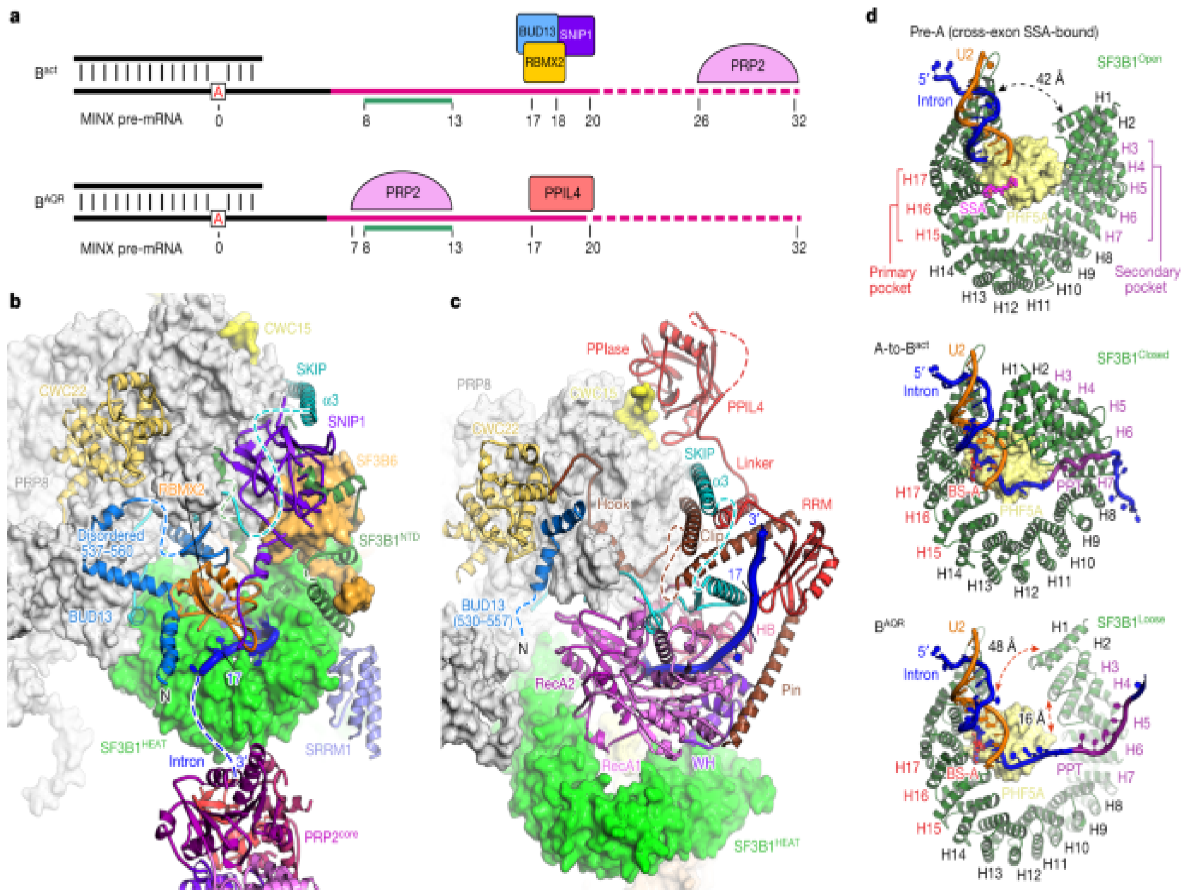
PRP2 translocates about 19 nucleotides

PRP2 has previously been observed at the periphery of human and yeast B^{act} complexes, residing on the convex side of the HEAT domain of SF3B1 (SF3B1^{HEAT})[5,11,13]. Notably, in the B^{AQR}, PRP2 is no longer anchored at the periphery of the spliceosome. Instead, PRP2 has now moved around 85 Å deep towards the branch helix and the core of B^{AQR}, filling a large cavity between SF3B1 and PRP8 within B^{AQR} (Fig. 2d). In the new location, PRP2 shares interfaces with SF3B1^{HEAT} and PRP8 through the RecA2 and HB domains, respectively (Extended Data Fig. 4b,c). The helicase rotated about 70° counterclockwise, and the RecA-like domains adopted the open conformation on top of SF3B1^{HEAT} (refs. [5,18]) (Fig. 2b,c). In this post-translocation state, PRP2 accommodated seven nucleotides of the intron,

assigned as the 7–13 region downstream of BS-A, in a channel framed by the RecA-like domains on one side and $PRP2^{CTD}$ on the other side.

Within yeast B^{act} complexes, PRP2 binds RNA at positions 28–34 (refs. [5,19]), which probably correspond to positions 26–32 in human B^{act} complexes according to our 3D alignment data (Extended Data Fig. 5a–d). A comparison between B^{AQR} and B^{act} indicated that PRP2 advanced approximately 19 nucleotides towards the spliceosome core during the B^{act}-to-B^{AQR} transition, which explains the stringent requirement for at least 32–34 nucleotides downstream of BS-A for efficient branching in human splicing[15] (Figs. 2d and 3a and Extended Data Fig. 5e). Therefore, PRP2 translocates along the intron from its earlier position in the B^{act} complex rather than acting from a distance like a molecular winch that pulls the intron from the periphery of the spliceosome (Supplementary Video 2). Concomitantly, the exiting intron was relocated about 70 Å owing to the helicase rotation and exchange in intron–protein interactions following the formation of the B^{AQR} complex (Fig. 2d and Extended Data Fig. 5f).

Fig. 3: PRP2 tightly regulates the dissociation of the RES complex and remodels the SF3B complex.



a, Relocation of PRP2 on the intron, RES complex dissociation from the translocated RNA and subsequent binding of PPIL4 to the latter. Pink, green and broken lines represent the PPT, the PPT region bound by SF3B1 within B^{act}, and the PPT-equivalent region visible only within the budding yeast B^{act} structure, respectively. **b**, Close-up view of the interfaces between PRP2 and other subunits in the B^{act} complex. **c**, The equivalent view and orientation of the B^{AQR} complex. **d**, SF3B1 exhibits primary and secondary hinged pockets for binding of BS-A and PPT, respectively. SF3B1 adopts a loose conformation in B^{AQR}, whereby the primary and secondary pockets are occupied and free, respectively.

Mechanism of dissociation of the RES complex

The RES complex is required to convert B to B^{act} complexes and dissociates after PRP2 action at the B^{act}-to-B* transition[20,21]. The RES complex comprises the proteins BUD13, RBMX2 (also known as SNU17) and SNIP1

(also known as PML1), and resides in a cavity of the B^{act} spliceosome framed by SF3B1 and PRP8. In this cavity, the RES complex binds the proteins SF3B1, PRP8, SKIP and CWC22, whereas the RRM domain of RBMX2 binds the polypyrimidine tract (PPT; or the equivalent region in yeast)[2,13].

In the B^{AQR} structure, the translocation of PRP2 displaces several subunits, including the RES complex from the top of $SF3B1^{HEAT}$ (Fig. 3b,c and Supplementary Video 3). The dissociation of RES is primarily caused by the stripping of RBMX2 away from the intron (Fig. 3a and Extended Data Fig. 5e,f). As BUD13 and SNIP1 share interfaces with RBMX2, dislocation of the latter probably releases the entire RES complex. Moreover, the clip of $PRP2^{NTD}$ replaced SNIP1 by interacting with the α3 helix of SKIP, whereas new interactions between $PRP2^{core}$, SKIP and PRP8 replace those between BUD13 and SKIP in the B^{act} complex (Fig. 3a and Extended Data Fig. 4c,d).

After destabilizing the RES complex from RNA and former protein contacts, only a short stretch of BUD13 (residues 530–557) is visible in B^{AQR}, which is bound to a composite structure formed by the hook of $PRP2^{NTD}$, PRP8 and the slightly relocated MA3 domain of CWC22 (Fig. 3a and Extended Data Fig. 4e,f). However, proteomics analysis showed that the entire RES complex remained bound to the spliceosome (Supplementary Data 1). Thus, BUD13(530–557) anchors the destabilized RES complex to the spliceosome, after being displaced from the intron. This flexible anchoring is conditioned by the hook of PRP2 binding to PRP8, the presence of which may serve as a signature for the occurrence of the translocation of PRP2. In this way, PRP2 might control the trajectory of the exiting RES complex, moving it away from the downstream intron region that must engage in new interactions.

SF3B1 unfastens the branch duplex

The remodelling and destabilization of the SF3B complex is a key event in the splicing pathway and a consequence of the translocation of PRP2. The largest subunit of SF3B is SF3B1, of which the HEAT domain exhibits a distinctive loose conformation in the B^{AQR} complex. This conformation

differs markedly from the open and closed conformations of SF3B1 from pre-A and from A-to-B^{act} complexes, respectively (Fig. 3d). For example, SF3B1(A514) and MINX(U136) are separated by 4, 48 and 42 Å in B^{act}, B^{AQR} and pre-A complexes, respectively. Furthermore, interacting residues SF3B1(H550) and PHF5A(Y51) from pre-A complexes are separated by 16 Å in the B^{AQR} complex.

The transition of SF3B1 from closed to loose following the conversion of B^{act} to B^{AQR} seemed to be caused by PRP2, after stripping the PPT (nucleotides 13–18) away from the binding pocket. Because the pocket is also a hinge, SF3B1 opens widely to unclamp the branch duplex from one side, like an opening pair of tweezers (Figs. 2b,c and 3d, Extended Data Fig. 6a,b and Supplementary Video 3).

The transition mainly involvs the H1–H10 repeats, whereas the contacts between BS-A and its binding pocket are almost identical in B^{act} and B^{AQR} complexes (Fig. 3d and Extended Data Fig. 6c). Overall, the consequence of the translocation of PRP2 is the disruption of 60% (around 1,100 $Å^2$) SF3B1's contact interface with RNA, facilitating the release of the branch duplex for relocation to the catalytic centre.

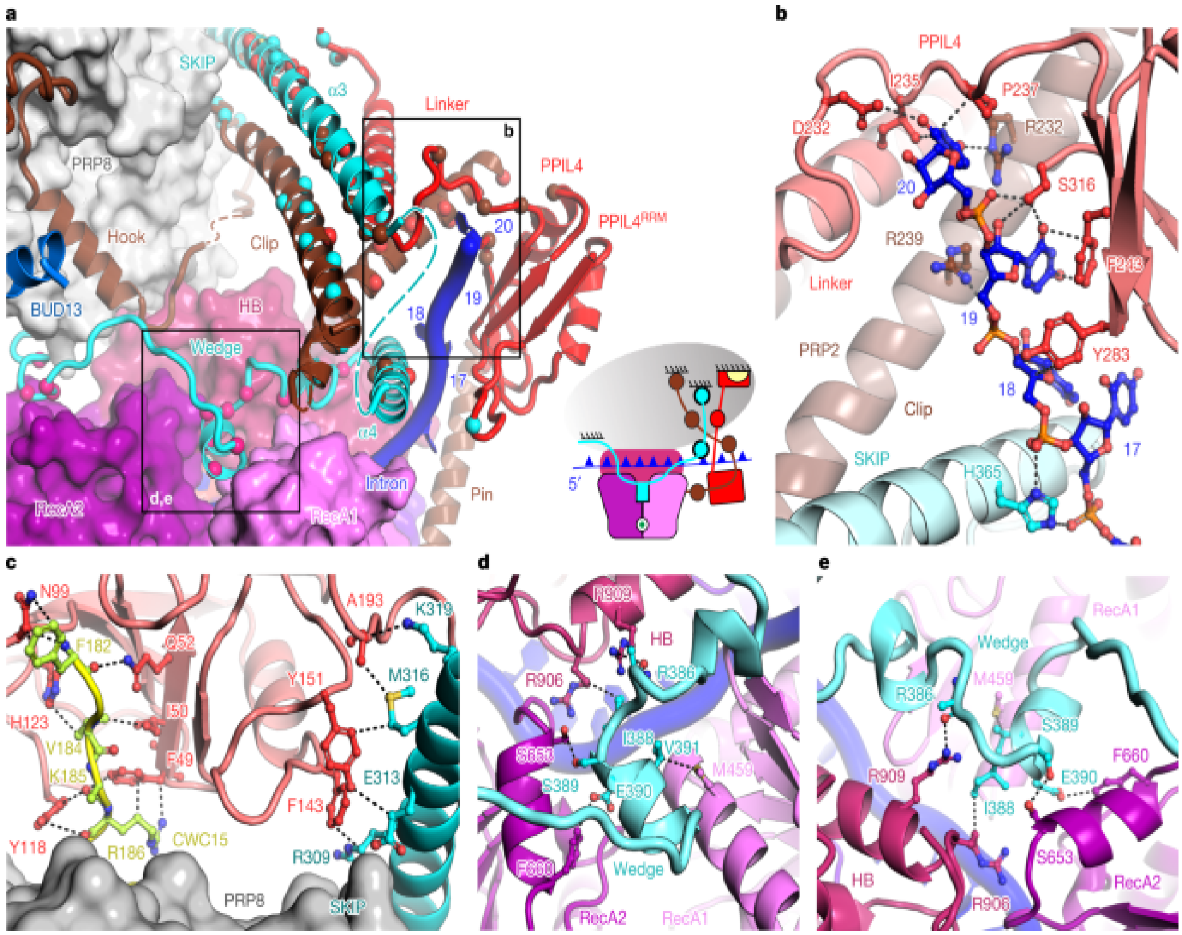
The opening of $SF3B1^{HEAT}$ by PRP2 triggers additional remodelling that weakened the interaction between the SF3B complex and the domains of PRP8 (Supplementary Video 3). In detail, the change in $SF3B1^{HEAT}$ curvature disrupts contacts between the repeats H10–H11 and $PRP8^{RH}$, and perturbs the endonuclease-like domain of PRP8 (Extended Data Fig. 6a). In addition, transition to the loose conformation induces the release of SRRM1 and SF3B6 from SF3B1. De-structuring of $SF3B1^{NTD}$ (residues 1–488) also occurs, which is an extended region that wraps around SF3B6 within the B^{act} complex (Fig. 3b,c and Supplementary Video 3). As SF3B6 and $SF3B1^{NTD}$ are connectors between the SF3B complex and PRP8 (Fig. 3c), their destabilization probably facilitates the subsequent release of SF3B and SF3A complexes, and the Aquarius-dependent relocation of the branch duplex at the transition from B^{AQR} to $B^{*}$and then C complexes (Supplementary Video 4).

Termination of the translocation of PRP2

The low complementarity between U2 and the branch sites of human introns requires tight regulation of the activity of PRP2 to prevent it from unwinding the branch duplex. Therefore, PRP2 needs to advance sufficiently far to dissociate the PPT from SF3B1 and open the SF3B1^{HEAT} clamp, yet not too far to compromise the branch duplex. Although PRP8 acts like an initial physical barrier able to stop the movement of PRP2 (Fig. 2d), the helicase may keep pulling the intron like a winch to unwind or alter the branch duplex. Thus, a different mechanism for the termination of PRP2 translocation is required.

PRP2 in B^{AQR} shares large interfaces with PPIL4 (1,040 Å^2) and SKIP (1,850 Å^2), which suggests that these proteins could regulate the helicase (Fig. 3c and Extended Data Fig. 4g–k). Indeed, flexible elements of PRP2^{NTD}, PPIL4 and SKIP form a stable 3D assembly that bound the exiting RNA strand at the nucleotides 17–20, downstream of the BS-A (Fig. 4a and Extended Data Fig. 4g,h). This result indicates that these proteins have a collective role in the timely termination of helicase translocation.

Fig. 4: PRP2 translocation terminates in a 3D construction reminiscent of a molecular brake.



a, Interactions between the subunits of the molecular brake. Interacting residues are shown as spheres depicted in the same colour as their binding partner. The domains of PRP2 RecA1, RecA2, HB and NTD are depicted. The anchoring surfaces of the brake elements to PRP8 (grey oval) are indicated. The inset shows a schematic representation of the molecular brake, in which the subunits are coloured as in the surface and cartoon representation. CWC15 is yellow. **b**, Interactions of the intron with PRP2^{NTD}, PPIL4 and SKIP. **c**, Interactions of PPIL4 with CWC15 and SKIP. **d**,**e**, The wedge element of SKIP intercalates between RecA1, RecA2 and the HB domains of PRP2, indicating inhibition of translocation. PPIL4, SKIP and CWC15 are red, cyan and yellow, respectively.

PPIL4 belongs to the cyclophilin family of peptidylprolyl isomerases (PPIases) and was not assigned in other cryo-EM structures of spliceosomes, which left its role in splicing unknown. In the B^{AQR} map, we built PPIL4, which included the PPIase and RRM domain separated by a linker

containing an α-helix (Extended Data Fig. 4g,h). Notably, $PPIL4^{RRM}$ and the linker bound the exiting intron, covering the position occupied previously by $RBMX2^{RRM}$ of the RES complex in B^{act} (Figs. 3a–c and 4a,b). Therefore, this intron-binding site is only accessible for PPIL4 after PRP2 has dissociated RES and translocated to an upstream position. Consistent with the B^{AQR} structure, recombinant PPIL4 formed a stable complex with PRP2 (137–1022) (delineated as the region built in the cryo-EM map; Extended Data Fig. 7a–c and Supplementary Fig. 2). Furthermore, PPIL4 was able to bind a model RNA substrate (for example, a poly(A)-rich 3′ single-stranded overhang) both in the presence and absence of PRP2 (Extended Data Fig. 7d,e and Supplementary Fig. 3).

Anchoring of PPIL4 to spliceosomes is facilitated by the NineTeen complex (NTC). The CWC15 component of NTC resides on the surface of PRP8, where it binds the $PPIL4^{PPIase}$ domain through a short stretch (Figs. 3c and 4c and Extended Data Fig. 4g). The defined distance between the PPIase and RRM domains of PPIL4 is maintained by the interdomains linker, structurally stabilized by helices from $PRP2^{NTD}$ and SKIP (Fig. 4a and Extended Data Fig. 4h).

SKIP is a largely unstructured protein, whose disparate elements interact with subunits of other spliceosomes[8]. The long characteristic α3 helix of SKIP (residues 282–340) binds PRP8 in B^{act}, B^{AQR} and C complexes (Fig. 3b,c and Extended Data Fig. 8a). In B^{AQR}, this helix stabilizes the post-translocation state of PRP2 through interactions with the clip of $PRP2^{NTD}$ and the linker of PPIL4. Furthermore, the residues 352–400 of SKIP are visible only in B^{AQR}, in which the α4 helix binds the clip of $PRP2^{NTD}$ and the intron at nucleotides 17–20 (Fig. 4a,b and Extended Data Fig. 4h). Notably, the residues 371–400 binds $PRP2^{core}$, such that a 3_{10} helix intercalates like a wedge between the two RecA modules and the HB domain of PRP2 (Fig. 4a and Extended Data Fig. 4j). Superposition of our structure with the PRP2 structure from *C. thermophilum* in the closed state (Protein Data Bank (PDB) identifier 6zm2)[18] indicates how the wedge element of SKIP might lock the helicase in the open conformation and therefore terminates translocation (Fig. 4d,e, Extended Data Fig. 4i and Supplementary Video 5). We suggest that PPIL4 and SKIP might act as

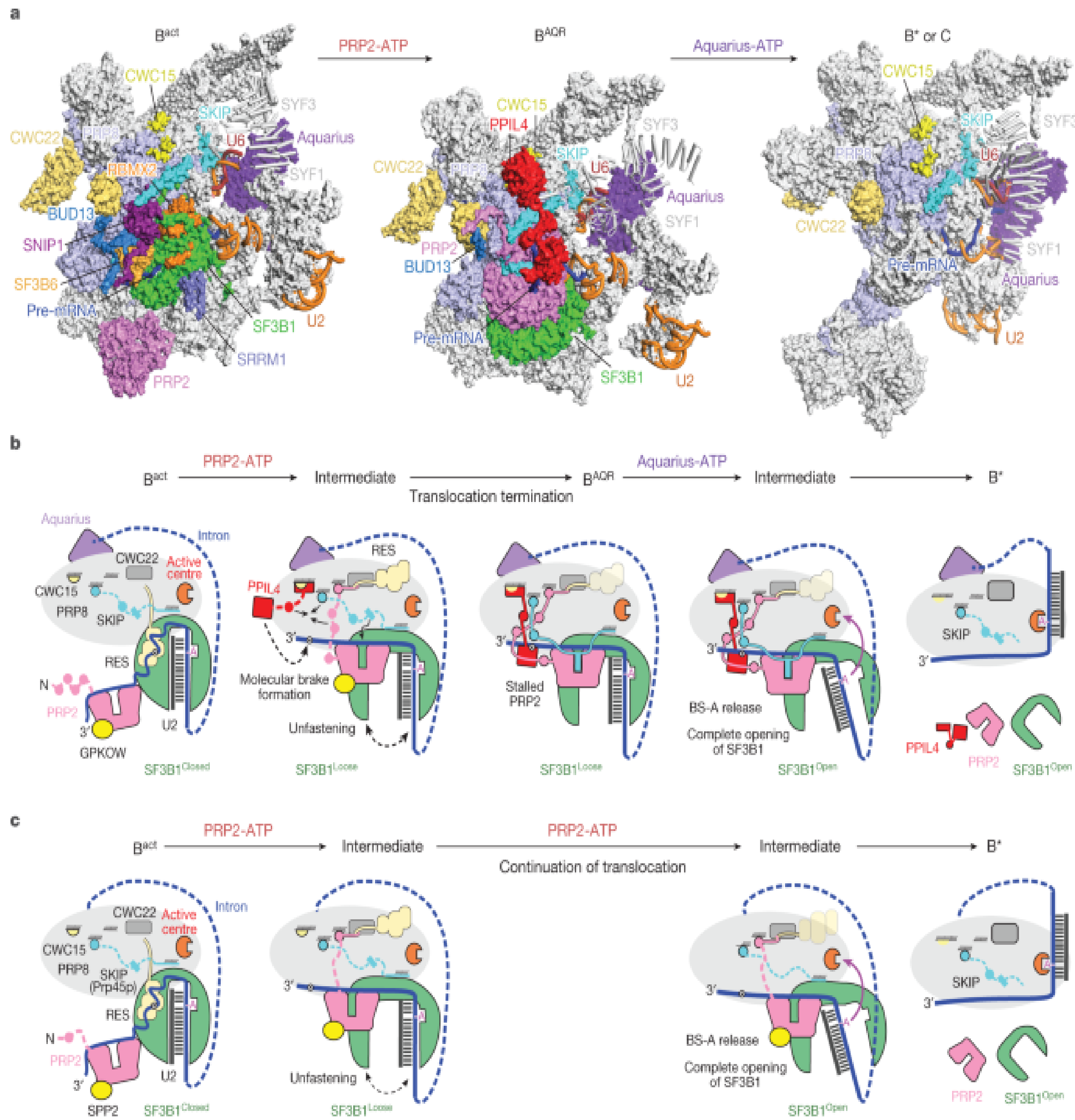
negative regulators of PRP2, forming a context-specific molecular brake to terminate the translocation of the helicase at a defined time and location on the intron.

Mode of action of a DEAH helicase

The four DEAH splicing helicases (PRP2, PRP16, PRP22 and PRP43) were observed at relatively fixed positions from the periphery of spliceosomes, where they act by less understood mechanisms[4]. In particular, whether PRP2 moves along the RNA or pulls the substrate from a distance remained unclear[3,4,19,22,23].

The structure of B^{AQR} reveals how PRP2 can operate in several stages (Fig. 5 and Supplementary Videos 2, 3 and 5). Starting from the periphery of B^{act} complexes[5,11,13], PRP2 translocates around 19 nucleotides to form the B^{AQR} complex while stripping away the proteins RBMX2 and SF3B1 from the intron. A cascade of events follows, including destabilizations (RES proteins), dissociations (SF3B6, SRRM1 and CWC27), restructuring ($SF3B1^{NTD}$, SKIP and $PRP2^{NTD}$), rearrangement of globular domains (PRP8 and CWC22) and recruitment (PPIL4). Essentially, the signal from ATP hydrolysis is amplified to induce large-scale remodelling.

Fig. 5: Model of catalytic activation in pre-mRNA splicing.



a, Catalytic activation depicted as structural transitions between B^{act}, B^{AQR} and B*. B* and the post-branching C complex have similar structures. Most subunits are shown as surface representations. SYF1 and SYF3 are shown as ribbons to enable better visualization of Aquarius. **b**, Structure-based model of human spliceosome remodelling by PRP2 and Aquarius. The hypothetical intermediate between B^{act} and B^{AQR} considers the following: (1) the molecular brake can form only after the binding of PPIL4 to the RNA, the binding site of which becomes available after dissociation of the RES complex by the translocation of PRP2; and (2) the advance of PRP2 should start stripping the intron from SF3B1 and promote the loose conformation of

SF3B1. The hypothetical intermediate between B^{AQR} and B* was generated by considering that Aquarius should induce complete displacement of the branch duplex from SF3B1, liberating BS-A from its binding pocket. Consequently, SF3B1 will transit from the loose to the open conformation. Because the HEAT repeats that bind PRP8 (in B^{AQR}) rearrange after loose-to-open conformation, SF3B1 dissociates from PRP8. This likely causes the complete destabilization of the SF3A–SF3B complexes and PRP2 from the spliceosome. **c**, Model of the spliceosome remodelling by Prp2p in budding yeast *(S. cerevisiae)*, based on the similarities and differences with the human counterpart.

The translocation of PRP2 terminates six nucleotides downstream of the BS-A in a network of interactions that resemble a molecular brake constructed from $PRP2^{NTD}$, SKIP and PPIL4. The brake binds the exiting intron like a brake shoe and intercalates a wedge element of SKIP between the RecA1, RecA2 and HB domains, immobilizing PRP2 in the open conformation (Fig. 4a,b,d and Supplementary Video 5). Thus, PPIL4 and SKIP might act as helicase cofactors from the class of negative regulators[24]. This termination mechanism differs from the constitutive inhibition of the DEAD-box helicase eIF4AIII in the exon–junction complex, whereby MAGO, Y14 and MLN51 appear to enclose and stabilize the RecA-like domains in the closed conformation[25]. Notably, CWC22 and NTC might contribute to the regulation of PRP2 by interacting with the RES complex and PPIL4, respectively. Finally, PRP2 dissociates during the spliceosome remodelling process driven by Aquarius.

Recombinant PRP2(137–1022) did not exhibit unwinding activity in vitro, either alone or in the presence of the cofactor GPKOW (Extended Data Fig. 7f–l and Supplementary Fig. 3). The same inactivity has been observed for the budding yeast counterpart Prp2p[26,27]. Whether this reflects a vital role of spliceosomes in activating translocation or an intrinsic inability of the helicase to use translocation for duplex separation remains unclear.

A question that is raised from these results is whether other splicing DEAH helicases also operate by translocation rather than winching[22,28,29]. PRP16 interacts with its targets in C complexes, the step I factors CCDC49 (Cwc25p in *S. cerevisiae*) and CCDC94 (Yju2 in *S. cerevisiae*), at the 3′ end

of the branched intron before their release[16,30]. Reminiscent of the function of PRP2, PRP16 might translocate in a 3′-to-5′ direction to displace the intron-bound step I factors, thereby promoting the repositioning of the pre-mRNA substrate for exon ligation. Furthermore, although introns bound by PRP22 are not visible in cryo-EM maps, PRP22 slightly relocates between different C* complex variants[31], which suggests that translocation is possible.

Identifying $PRP2^{NTD}$ as an accessory domain for multiple context-specific functions might provide suggestions on how NTDs function in other DEAH helicases. For instance, a fragment of $PRP22^{NTD}$ (residues 393–427) inserts between the RT and linker domains of PRP8, which indicates that they have a role in anchoring the helicase downstream of the 3′ exon in the human C* and P spliceosomes[31,32,33]. Overall, the mechanism of action of PRP2 captured within B^{AQR} may provide a paradigm for other DEAH helicases and guide future investigations.

Two helicases drive catalytic activation

An important implication of this work is that the B^{act} to B* transition, also known as catalytic activation, is driven by two helicases—PRP2 and Aquarius, acting on the branch site from downstream and upstream locations, respectively (Figs. 1b and 5b).

Arrested between the action of the two helicases, the B^{AQR} structure revealed how PRP2 operates at the B^{act}-to-B^{AQR} transition and implicates Aquarius as the helicase that promotes the conversion of B^{AQR} to B^{*}. The requirement for two helicases is probably an adaptation for transferring the delicate branch duplex—which is often unstable owing to low complementarity[34,35]—to the catalytic centre while keeping it unaltered. Therefore, whereas PRP2 operates downstream of the branch site to unfasten the branch duplex from the constraints of the SF3B–RES complex, Aquarius acts upstream of the branch site to induce the removal of the remaining constraints and the relocation to the catalytic centre.

This two-step catalytic activation process could provide additional checkpoints to proofread and discard incorrectly selected introns, the length and sequence of which can vary substantially in metazoans[36,37]. This mechanism might also facilitate coordination between splicing and other intron-related events. Of relevance, Aquarius couples splicing to the biogenesis of intron-encoded box C/D small nucleolar RNP (snoRNP)[38], which suggests that it coordinates separation in time and space between snoRNP biogenesis and branch duplex transfer for splicing.

Aquarius binds the intron at position −40 to −33 upstream of BS-A in C complexes[38]. Despite the moderate resolution around Aquarius (around 6.2 Å), a comparison between B^{AQR} and C complexes showed a repositioning of the helicase, which suggests how it might operate (Extended Data Fig. 9). We have previously shown that Aquarius unwinds duplexes in vitro with apparent 3′-to-5′ polarity. As the mutant Aquarius(Y1196A) inactivates unwinding but not splicing, the relevance of this assay on splicing remains unclear[6]. However, putative translocation with 3′-to-5′ polarity might induce relaxation of the intron and perturb the contacts between the domains of Aquarius and adjacent subunits. This movement could propagate changes in interfaces between proteins interposed from Aquarius to the SF3B1–BS-A interface (primarily SF3A and SF3B subunits), which causes the release of BS-A and dissociation of the SF3A and SF3B complexes (Extended Data Fig. 9a–c).

Aquarius might also use 5′-to-3′ polarity, similar to the phylogenetically related SF1 helicase Upf1. By pulling the intron with this polarity, Aquarius could remotely eject BS-A from the primary pocket of SF3B1. Consequently, SF3B1 would switch from a loose to an open conformation, thereby restructuring HEAT repeats 16–20, which forms interfaces with PRP8, SF3B2 and SF3A2 (Extended Data Fig. 9d–g). Ultimately, the entire SF3A–SF3B complex would undock from the core components PRP8 and the snRNA U6, which accomplishes the conversion of B^{AQR} to B* and C complexes (Fig. 5b, Extended Data Fig. 8 and Supplementary Video 4).

The arrest of PRP2 in interactions with the molecular brake indicates that Aquarius starts acting after the complete translocation of PRP2. However, some chronological overlap of the actions of the two helicases cannot be

excluded. For instance, Aquarius might change the tension in the intron to induce additional loosening of SF3B1 while PRP2 facilitates the complete liberation of the branch duplex.

Significantly, Aquarius and all molecular brake components—PPIL4, elements of PRP2^{NTD} and SKIP—are highly conserved among humans, plants and yeasts[39] (Supplementary Fig. 4). This conservation supports the view that Aquarius and the molecular brake have coevolved to complement PRP2 across all kingdoms of eukaryotes, which indicates that catalytic activation in two ATP-dependent steps might be a universal mechanism.

Simplified catalytic activation in yeast

In contrast to fission yeast (*Schizosaccharomyces pombe*), budding yeast (*S. cerevisiae*) lacks Aquarius and the molecular brake, which raises the question of how catalytic activation can occur only by PRP2 (Prp2p in *S. cerevisiae*). Based on this work and a large body of published data about Prp2p, we propose a model for spliceosome remodelling in budding yeast (Fig. 5c).

Prp2p binds the 3′ tail of the intron, advancing in the 3′-to-5′ direction to disrupt interactions between SF3b1 (also known as Hsh155p) and the branch duplex while destabilizing SF3a, SF3b and RES complexes[19,23,40,41]. Consistent with the human model, this remodelling implies that Prp2p uses translocation rather than winching, as postulated in earlier models[19,40] (Fig. 5c). Spp2p, the G-patch orthologue of human GPKOW, assists these remodelling steps[1,5,27].

After destabilization of Res, we propose that Prp2p might transit into a B^{AQR}-like state, arriving at approximately 6 nucleotides downstream of the branch duplex. This scenario could explain the genetic interaction between the cold-sensitive allele *Prp2p*^{Q548N} (*PRP2*^{Q721} in humans) and the SF3b1(D450G) or SF3b1(V502F) mutant[42,43]. The human counterparts SF3B1(D781) and SF3B1(L833) are in physical proximity to PRP2 only within B^{AQR} spliceosomes (Extended Data Fig. 4l).

Furthermore, the human B^{AQR} structure might explain the effect of Cwc22p(454–491) on the ability of Prp2p to remodel the B^{act} spliceosome[44]. The equivalent region of CWC22, the hook of $PRP2^{NTD}$ (conserved in budding yeast) and the RES complex interact in B^{AQR}, which suggests that the role of Cwc22p is conserved in facilitating the function of the helicase[44] (Fig. 3b and Extended Data Fig. 4m–o). Notably, $Prp2p^{NTD}$ is dispensable for viability[45], which suggests that neighbouring subunits might functionally compensate for its deletion.

In contrast to the human orthologue, Prp2p might advance beyond the B^{AQR}-equivalent time point, as there is no molecular brake to terminate translocation. The unrestricted Prp2p could extract BS-A from the primary pocket, which triggers the transition of SF3b1 from the loose to the close conformation. In this way, Prp2p liberates the branch duplex and dissociates SF3a, SF3b and RES complexes, and the proteins Cwc24p and Cwc27p[19,23,41] (Fig. 5c). We suggest that this simplified mechanism is tolerated in budding yeast primarily because the high complementarity between the intron and U2 snRNA prevents its alteration by the unidirectional action of Prp2p.

The complete cycle of SF3B1 transitions

Accumulating evidence outlines SF3B1 as a conformational switch adapted for the recognition and transfer of branch sites. This work provides new insight into the function and cyclic pathway of conformational transitions of SF3B1, mediated by RNA helicases (Extended Data Fig. 6d).

First, the ability of SF3B1 to switch conformations relies on two pockets—here referred to as primary and secondary pockets—that sequentially bind BS-A and PPT after the formation of the prespliceosome or release them in the reverse order during catalytic activation (Fig. 3d and Extended Data Fig. 6d). Both pockets are also conformational hinges, which thereby induce the progressive compaction or decompaction of the HEAT domain following the binding or release of cognate RNA elements. Consequently, SF3B1 transits stepwise to the open, half-closed and closed conformation after BS-A and PPT recognition during pre-A conversion to A complexes[46,47,48]. The open

state facilitates the formation of the branch duplex by toehold-mediated strand invasion—a non-enzymatic mechanism captured in a human pre-A complex stalled using spliceostatin A[47].

The closed conformation persists until the B^{act} complex, when PRP2 detaches the bound PPT from the secondary pocket, which then induces the loose conformation of SF3B1 within B^{AQR}. The requirement of bound PPT for the closed conformation might explain the formation of B^{act}-like spliceosomes on substrates bearing at least 16 nucleotides downstream of BS-A, whereas 6 nucleotides are insufficient[15]. Notably, the loose conformation differs from the half-closed one, which is probably due to the contacts between the $PRP2^{core}$ and HEAT repeats (410 $Å^2$ interfaces; Fig. 2b,c).

Next, Aquarius promotes the extraction of BS-A from the primary pocket and the release of SF3B, probably in the open conformation typical for the apo form[46]. The released SF3B might be recycled as a building block for the biogenesis of 17S U2 snRNPs assisted by the helicase SF3b125 (ref. [49]), thus re-entering the splicing pathway. Overall, the structure of B^{AQR} reveals the crucial importance of SF3B1 as a conformational switch in catalytic activation and how helicases can modulate the pathway of conformational transitions of SF3B1.

Methods

Cloning

Recombinant proteins were produced from codon-optimized synthetic genes (GeneArt, ThermoFisher Scientific). Full-length wild-type Aquarius and its K829A mutant were fused to a C-terminal 8×His-tag and cloned into the pFL vector backbone as previously described[6]. For co-expression of IBC subunits, the master, 4-protein construct ΔISCC, comprising SYF1 (also known as XAB2), ISY1, CCDC16 (also known as ZNF830) and PPIE (also known as CypE), was assembled by in vitro Cre–*loxP* recombination of the donor pSPL-CCDC16–PPIE and acceptor pFL-SYF1–ISY1 vectors[6]. The expression constructs were transformed into DH10MultiBacY *Escherichia*

coli cells. The resulting bacmids were isolated using a High Pure Plasmid Isolation kit (Roche) and used for transfection into Sf9 (*Spodoptera frugiperda*) insect cells. Sf9 cells were transfected with the help of either X-tremeGENE HP (Roche) or FuGENE HD (Promega) reagents. Insect cell lines were not tested for mycoplasma contamination.

Human PRP2 (NCBI Reference Sequence identifier: NM_003587, transcript variant 1) and human PPIL4 (NCBI Sequence identifiers BC020986 and NM_139126.4) open reading frame clones were obtained from OriGene (RC202912) and Applied Biological Materials (373710120000), whereas human GPKOW and SKIP were codon-optimized for expression in insect cells and synthesized by GeneArt (ThermoFisher Scientific). All constructs were inserted by ligation-independent cloning into a modified pFastBac vector backbone in-frame with an N-terminal twin-StrepII affinity tag, which can be cleaved off with the HRV-3C protease. All recombinant constructs were verified by Sanger sequencing.

Purification of recombinant proteins

Recombinant proteins were produced in insect cells using the MultiBac system[50]. The initial V_0 and V_1 baculovirus stocks were produced in Sf9 cells. Large-scale protein co-expression was conducted in High Five (Hi5) insect cells (*Trichoplusia ni*, BTI-TN5B1-4) through the combination of two V_1 stocks of baculovirus, coding for the C-terminal 8×His-tagged Aquarius (wild-type or mutant K829A) and the untagged ΔISCC complex, in 1:1.5 ratio (V_1^{AQR}:$V_1^{\Delta ISCC}$). The purification of the wild-type and the K829A dominant-negative IBCs was performed according to a previous protocol established in our laboratory, but with some modifications[6]. Insect cells expressing IBC were disrupted by sonication in lysis buffer (50 mM HEPES-NaOH, pH 7.5, 400 mM NaCl, 10% (v/v) glycerol, 5 mM 2-mercaptoethanol (2-ME), 30 mM imidazole and protease inhibitors (complete EDTA-free, Roche)) and the cell debris was pelleted by centrifugation at 18,000 r.p.m. at 4 °C for 1 h (A27-8×50 rotor; Thermo Scientific). The supernatant was loaded on a 5 ml HisTrap HP Ni(II)-chelating column (GE Healthcare/Cytiva) equilibrated with the lysis buffer. The Ni(II)-bound protein species were eluted using a linear 30–400 mM imidazole gradient in Ni(II) elution buffer (50 mM HEPES-NaOH, pH 7.5,

200 mM NaCl, 10% (v/v) glycerol and 5 mM 2-ME) and analysed by SDS–PAGE. The IBC-containing fractions were loaded on an anion-exchange Q Sepharose HP column (GE Healthcare/Cytiva) equilibrated in 20 mM HEPES-KOH, pH 7.5, 200 mM KCl, 10% (v/v) glycerol and 5 mM dithiothreitol (DTT). The bound IBC was eluted off the resin with a linear 0.2–1 M KCl gradient. The peak fractions, corresponding to the stoichiometric IBC, were pooled, concentrated to 10 mg ml^{-1}, aliquoted, flash-frozen in liquid nitrogen and stored at −80 °C.

Twin-StrepII-tagged human PRP2(137–1022), PPIL4 and GPKOW were expressed in Sf9 or Hi5 insect cells using recombinant baculoviruses, prepared as described above, in amounts sufficient to induce cell cycle arrest in 24 h. All recombinant proteins used in functional assays were purified at 4 °C using similar purification protocols. In brief, insect cells, expressing the target proteins, were briefly sonicated and lysed in lysis buffer (50 mM HEPES-KOH, pH 7.5, 150 mM KCl, 10% (v/v) glycerol, 0.2% (v/v) Triton X-100 RNase-free, 2 mM DTT and protease inhibitors (complete EDTA-free, Roche)), with the detergent added after sonication, and the cell debris was pelleted by centrifugation at 23,000 r.p.m. for 1 h in an A27-8×50 rotor (Thermo Scientific) or 40,000 r.p.m. (approximately 164,244*g*) in a type 70 Ti rotor (Beckman Coulter) for 45 min at 4 °C. The filtered supernatant was applied to a 5 ml Strep-Tactin XT 4Flow pre-packed column (IBA Lifesciences) or incubated in batch for 1 h with around 1 ml or 2.5 ml Strep-Tactin beads per 1 litre of Sf9 or Hi5 culture, respectively. The bound proteins were eluted by competition with 50 mM biotin in lysis buffer or with elution buffer containing 60 mM biotin (50 mM HEPES-KOH, pH 7.5, 100 mM KCl, 5% (v/v) glycerol, 1 mM EDTA, 2 mM DTT and 60 mM biotin). The Strep-Tactin eluates were concentrated and further purified by size-exclusion chromatography (SEC) on Superdex 200 HiLoad 16/600 200 pg (GE Healthcare/Cytiva) or Superdex 200 Increase 10/300 GL (GE Healthcare/Cytiva) equilibrated in 20 mM HEPES-KOH, pH 7.5, 150 mM KCl, 10% (v/v) glycerol and 2 mM DTT. Alternatively, the PRP2(137–1022) Strep-Tactin eluates were purified on Q Sepharose HP (GE Healthcare/Cytiva) and eluted from the 5 ml anion-exchange column using a linear 0–30% gradient formed over 40 ml between buffer A (20 mM HEPES-KOH, pH 7.5, 100 mM KCl, 5% (v/v) glycerol and 0.5 mM DTT) and buffer B (20 mM HEPES-KOH, pH 7.5, 1 M KCl, 5% (v/v) glycerol and 0.5 mM

DTT). With the exception of PPIL4, all proteins used in functional assays comprised the N-terminal Twin-StrepII tag. The Twin-StrepII tag of PPIL4 was removed by on-bead cleavage with the HRV-3C protease, and the untagged protein was purified by SEC on Superdex 200 Increase 10/300 GL (GE Healthcare/Cytiva). Purified proteins were concentrated by ultrafiltration to their stock concentration (PRP2(137–1022), around 19–22 µM; GPKOW, around 165 µM; untagged PPIL4, around 47 µM; tagged PPIL4, about 67 µM), frozen in liquid nitrogen and stored at −80 °C or used directly in assays.

Reconstitution of PRP2–GPKOW and PRP2–PPIL4 complexes

Using a minimal in vitro system, we first tested direct interactions and stable formation of complexes between PRP2(137–1022) and its cofactors GPKOW and PPIL4 as subjected to SEC. All SEC analyses were carried out in SEC buffer (20 mM HEPES-KOH, pH 7.5, 100 mM KCl, 1.5 mM $MgCl_2$ and 5% (v/v) glycerol). The PRP2 complexes were reconstituted in vitro by mixing around 30 µg recombinant helicase with a 5-fold excess of cofactor (PPIL4 or GPKOW) in a 100 µl reaction volume. The samples were then incubated for 1 h on ice and applied to a Superdex 200 Increase 10/300 GL column (GE Healthcare/Cytiva) run at 0.5 ml min^{-1} using an Äkta Go system (Cytiva). SDS–PAGE analyses of the SEC fractions showed that in both cases, the peak profile of PRP2(137–1022) shifted to early fractions in the presence of a cofactor (fraction 9 in the presence of PPIL4 and fraction 7 in the presence of GPKOW) compared with the protein alone (fraction 11), which indicated the formation of a helicase–cofactor complex. As independent means of probing the direct interaction of PRP2(137–1022) with PPIL4, we co-expressed the two splicing factors in Sf9 insect cells and affinity purified their complex from baculovirus-infected cultures. The cultured cells were lysed in lysis buffer (50 mM HEPES-KOH, pH 7.5, 150 mM KCl, 5% (v/v) glycerol, 2 mM DTT, 2 mM $MgCl_2$, 0.1% (v/v) Triton X-100 and protease inhibitors (complete EDTA-free, Roche)) and the expressed factors were captured on Strep-Tactin XT 4FLOW affinity beads (IBA Lifesciences). The protein samples were eluted by competition with biotin in elution buffer (50 mM HEPES-KOH, pH 7.5, 150 mM KCl, 5% (v/v) glycerol, 2 mM DTT, 2 mM $MgCl_2$ and 60 mM biotin), concentrated by ultrafiltration to about 500 µl and subjected to SEC in SEC buffer. As for

the in vitro assembled PRP2(137–1022)–PPIL4 complex, the complex prepared from insect cells peaked in fraction 9 when eluting from the Superdex 200 column. Overall, this result shows that the limited interaction interface between the NTD of PRP2 and PPIL4, observed in the B^{AQR} cryo-EM structure, promotes the stable recruitment of the helicase cofactor in a RNA-independent manner. The PRP2(137–1022)–PPIL4 complex used in helicase activity assays was prepared by insect cell co-expression, purified in two steps, as described above, and concentrated to about 14.5 μM. The PRP2(137–1022)–GPKOW complex used in biochemical assays was assembled in vitro, purified by SEC from GPKOW excess and concentrated to about 40.3 μM. In vitro reconstitution of the complex between PRP2(137–1022) and PPIL4 was performed two times using two independent protein preparations. The PRP2(137–1022)–GPKOW complex was assembled in vitro at least three times.

Preparation of the nuclear extracts active in splicing

HeLa S3 cells, tested for mycoplasma, were obtained from GBF (Helmholtz Centre for Infection Research). The nuclear extract active in splicing was prepared according to the standard protocol from Dignam and used as described[13]. Cells were grown in a 30 litre fermenter (Applikon Biotek) to a density of 6.5×10^6 cells per ml in DMEM/F12 (1:1) medium supplemented with 5% (v/v) newborn calf serum. After collection by centrifugation for 10 min at 2,000 r.p.m. in an 8 × 2,000 ml BIOS rotor (Thermo Scientific), the cells were washed twice with cold 1× PBS buffer. The cell pellet was re-suspended in MC buffer (10 mM HEPES-KOH, pH 7.6, 10 mM potassium acetate, 0.5 mM magnesium acetate, 0.5 mM DTT and 2 tablets of protease inhibitors (complete EDTA-free, Roche) per 50 ml of the buffer). After 5 min of incubation on ice, the cells were lysed with 18 strokes of a Dounce homogenizer at 4 °C. The nuclei were pelleted for 5 min at 10,000 r.p.m. in a F14-14×50cy rotor (Thermo Scientific) and were further lysed in Roeder C buffer (20 mM HEPES-KOH, pH 7.9, 0.2 mM EDTA, pH 8.0, 25% (v/v) glycerol, 420 mM NaCl, 1 mM $MgCl_2$, 0.5 mM DTT and 0.5 mM PMSF) by 20 strokes of a Dounce homogenizer at 4 °C. The mixture was stirred slowly for 40 min at 4 °C and centrifuged at 12,300 r.p.m. in a F14-14×50cy rotor. The supernatant corresponding to the active nuclear extract was aliquoted and flash-frozen in liquid nitrogen and stored at −80 °C.

In vitro pre-mRNA splicing

A typical splicing reaction contained the following components: 20% (v/v) non-dialysed HeLa nuclear extract, 3 mM $MgCl_2$, 2 mM ATP, 20 mM creatine phosphate and 10 nM m^7G-capped MINX-3×MS2 pre-mRNA, fluorescently body-labelled with cyanine 5-uridine-5′-triphosphate (Cy5-UTP, Enzo). To monitor splicing of the Cy5-labelled pre-mRNA substrate, the splicing reactions were incubated for 15, 30, 60, 90 and 120 min at 30 °C (Extended Data Fig. 1c). To assess the effect of recombinant IBC(K829A) on the assembly of the spliceosome, the splicing reaction was supplemented with IBC(K829A) to a final concentration of 0.45 μM and pre-incubated for 20 min at 30 °C before the addition of the pre-mRNA substrate and ATP (Extended Data Fig. 1c). In all cases, RNA was recovered by phenol–chloroform–isoamyl alcohol extraction followed by ethanol precipitation. The recovered RNA samples were analysed on denaturing urea polyacrylamide gels (14%, 0.5× TBE). The fluorescently labelled pre-mRNA substrate, splicing intermediates and the final products were detected by in-gel fluorescence using a Typhoon FLA 9500 imaging system (GE Healthcare).

Affinity purification of the human B^{AQR} spliceosome

Spliceosomal complexes stalled with recombinant IBC(K829A) were assembled in vitro on non-labelled m^7G-capped MINX-3×MS2 pre-mRNA, synthesized by T7 RNA polymerase run-off transcription. Before initiating the reaction, pre-mRNA substrate (used at 10 nM, final concentration) was incubated with a 20-fold molar access of MBP-MS2 fusion protein for 30 min at 4 °C. Before the addition of ATP and pre-mRNA-MBP-MS2, the large-scale splicing reactions were supplemented with 0.45 μM recombinant IBC(K829A) complex and incubated for 20 min at 30 °C. Therefore, a typical preparative splicing reaction, used from cryo-EM sample preparation, comprised (final concentrations) the following: 20 mM HEPES-KOH, pH 7.9, 2 mM ATP, 20 mM creatine phosphate, 3.2 mM $MgCl_2$, 84 mM NaCl, 5% (v/v) glycerol, 20% (v/v) non-dialysed HeLa nuclear extract, 10 nM MINX-3×MS2, 200 nM MBP-MS2 and 40 μM IBC(K829A).

To enable the spliceosome assembly, the splicing mixture was slowly stirred for 120 min at 30 °C in a water bath. The non-incorporated pre-mRNA was cleaved by DNA-directed endogenous RNase H digestion for 20 min at 30 °C using a 30-fold molar excess of the cmd42 (5′-TCTTACCGTTCG-3′) and cmd43 (5′-CGGGTTTCCGAT-3′) antisense DNA oligonucleotides. To prevent precipitation of assembled spliceosomes, NaCl was slowly added to the splicing reaction to a final concentration of 120 mM. The aggregates were removed by centrifugation for 10 min at 12,300 r.p.m. in a F14-14×50 rotor. The supernatant was then applied onto a pre-packed 5 ml MBPTrap HP column (GE Healthcare/Cytiva) at 1 ml min^{-1}, washed with 20× column volumes of G120 buffer (20 mM HEPES-KOH, pH 7.9, 1.5 mM $MgCl_2$ and 120 mM KCl) and the spliceosomal complexes were eluted with G120 buffer supplemented with 3 mM maltose. The peak fractions were analysed by denaturing 4–12% NuPAGE gels (Life Technologies). The gels were stained with SYBR Gold (Invitrogen) and Coomassie to detect RNA and protein species, respectively. The elution fractions containing IBC(K829A) spliceosomes (B^{AQR}) were loaded onto a linear 10–30% (w/v) glycerol gradient prepared in G120 buffer and centrifuged at 20,500 r.p.m. for 16 h at 4 °C in a TST 41.14 rotor (Kontron). The gradients were divided into 23 fractions (500 µl each) and manually collected from the top. Peak fractions (14, 15 and 16) from 3 different gradients were pooled and crosslinked with 0.1% (v/v) glutaraldehyde (Electron Microscopy Sciences) for 1 h on ice. The unreacted glutaraldehyde was quenched with 100 mM aspartate (final concentration). The crosslinked B^{AQR} was then incubated for 2 h on ice and concentrated to the sample absorbance at 280 nm of 0.6. The buffer was exchanged by ultrafiltration to the sample buffer (20 mM HEPES-KOH, pH 7.9, 120 mM NaCl, 1.5 mM magnesium acetate and 2% (w/v) glycerol) using an Amicon 50 kDa MWCO (Millipore) spin concentrator and used directly for cryo-EM grid preparation.

Preparation of cryo-EM grids

Volumes of 4 µl of concentrated sample were applied to one side of glow-discharged UltrAuFoil 200 R2/2 grids (Quantifoil) in a Vitrobot Mark IV (FEI) operated at 4 °C and 100% humidity. The grids were blotted for 2 s with a blotting force of 5 and immediately frozen by plunging into liquid ethane cooled by liquid nitrogen.

Cryo-EM data acquisition and image analysis

The electron micrographs for all datasets were acquired on an FEI Titan Krios G2 transmission electron microscope operated at 300 keV in EFTEM mode, equipped with a Quantum LS 967 energy filter (Gatan), zero loss mode, 30-eV slit width and a K2 Summit direct electron detector (Gatan) in counting mode. Automated data acquisition for dataset 1 (untilted) and dataset 2 (tilted, 25°) was performed using EPU software (Thermo Fisher) at a nominal magnification of ×130,000 (1.05 Å per pixel (Å px^{-1})). Micrographs for these two datasets were collected as 40-frame movies at a dose rate of around 5 e^{-} $Å^{-2}$ s^{-1} over 9 s, which resulted in a total dose of about 45 e^{-} $Å^{-2}$.

The two cryo-EM datasets of B^{AQR}, collected at 0° and 25° tilt angles, were separately preprocessed on the fly with Warp (motion correction, CTF estimation and dose weighting), and particles were picked using a retrained convolutional neural network[51]. Each set of Warp-extracted particles was then subjected to three parallel 2D classification runs in cryoSPARC (v.0.65) using 50 classes and applying a class uncertainty factor of 1.5. The good classes were selected from the three runs, and all particles in these good classes were merged (while removing duplicates).

Initial attempts at processing the cryo-EM images in cryoSPARC revealed that this spliceosome complex exhibited a more structurally rigid core, which continued with more dynamic and poorly resolved peripheral regions (Extended Data Fig. 2). Hence, we systematically tested different EM data processing routines with the aim of improving the more peripheral densities of the spliceosome and to computationally resolve its structural heterogeneity. In brief, the Warp-picked, combined B^{AQR} particle images (734,691 particles) were re-extracted in Relion 3.1 using a box size of 640/640 px (672 Å/672 Å) and then 2× binned before being subjected to 2D classification with the 'Ignore CTFs until the first peak' option switched on (Extended Data Fig. 2). The resulting 'good' particles subset (160,650 particles) was then refined in 3D using a 60 Å low-passed human B^{act} map as a reference volume (Electron Microscopy Databank identifier EMD-4236)[13] and a spherical mask. The obtained consensus map of the complex (map M1) was then used as a reference in a global 3D classification round

with 10 classes starting from the initial pool of B^{AQR} particles (Extended Data Fig. 2). The cleaned subset of B^{AQR} particles was then subjected to another round of global 3D classification with 8 classes and a 20 Å resolution limit. This second-round of 3D classification enabled us to resolve the two major compositional states of the complex (Extended Data Fig. 2), termed state A (29.2% particles) and state B (3.4% particles). In the state B complex, stronger density was observed for the PRP19 helical bundle and the step two splicing factor PRP17. Conversely, these map regions appeared to be poorly resolved in the state A complex, which is probably due to their flexibility and low occupancy, as observed for a previously described B^{act} complex[13]. In both cases, however, the peripheral building blocks of the complex (the IBC, U2 3′ core and the BRR2 helicase) were less resolved than the central part of the complex. To further improve the density of the state A map, we performed an additional global 3D classification without image alignment with 8 classes and a 20 Å resolution limit and then re-extracted and re-centred the B^{AQR} particle images at their original sampling rate (1.05 Å px^{-1}) in a 520 px box. The 3D refinement of this subset of particles (179,552 particles) with a soft mask was then followed by CTF refinement (per particle defocus, per micrograph astigmatism) and an additional round of classification in 2D. The subsequent refinement in 3D of this final subset of particles (146,157 particles) with loose or tight soft masks around B^{AQR} resulted in map 2 and map 3 of the complex that reached global, gold-standard Fourier shell correlation resolutions of about 3.1 Å and 2.9 Å, respectively (Extended Data Figs. 2 and 3). Using these well-resolved consensus maps, we modelled the translocating PRP2 helicase, including its extended NTD domain, as well as the remodelled SF3B complex, and two out of the four B^{AQR} PPIases (that is, PPIL2 and PPIL4).

Starting from the particle sets, which led to the higher resolution core maps, we also carried out local 3D classifications (Extended Data Fig. 4). Thus, by applying a mask covering the overall B^{AQR} volume, we subjected these particles to 3D classification without image alignment with 6 classes. Particles exhibiting more pronounced densities for the U2 3′ core module and the IBC module (27,719 particles) were subsequently re-extracted and refined in 3D with soft and spherical masks. These alternative B^{AQR} maps enabled the rigid body placement of the U2 3′ core and IBC modules at the

periphery of B^{AQR}. Local classifications were also performed on the particle set assigned to state B of the complex. In this case, a soft mask was applied to the region of the spliceosome where the BRR2 helicase resides, which was poorly resolved in the consensus map. The state B particle images were then subjected to 3D classification without image alignment and 4 classes. A minority subset of the overall particle set (12,395 particles) appeared to contain a more ordered BRR2, with the helicase being generally destabilized in B^{AQR} cryo-EM images, which is probably because of the substantial change to the conformation of SF3B1 induced by the translocation of PRP2. These particles were then re-extracted and refined in 3D to obtain the complete map of the complex (map M4), which now enclosed the U2 3′ core, the IBC module, the PRP19 helical bundle, the BRR2 helicase and the splicing factor PRP17.

Density assignment, model building and refinement

To enable model building, the B^{AQR} core maps (Supplementary Fig. 2) were sharpened using DeepEMhancer[52] or locally scaled with LocScale[53]. The quality of the cryo-EM density map at the core of B^{AQR} enabled careful model building and side-chain assignment, whereas modelling of the more solvent-exposed map regions was restricted to backbone tracing and rigid-body docking of known structures and computational models. Initial interpretation of the B^{AQR} maps was facilitated by the available models of human B^{act} complexes obtained in several states (PDB identifiers 6FF4, 6FF7 and 5Z57). Thus, the model building of the B^{AQR} complex was initiated by the docking of the human B^{act} models into the consensus maps (maps 2, 3 and 4) of the complex, followed by manual, residue-by-residue model adjustment in Coot[54] and refinement with phenix.real_space_refine[55]. Although filtered, locally scaled maps were used for model interpretation, the real space refinement of the B^{AQR} model was carried out exclusively against the original, unsharpened maps. The higher local resolution of our maps enabled us to improve and correct some of the available models for the B^{act} core region. After initial placement and refinement, several new (or reconfigured) B^{AQR} density regions were observed compared with previous B^{act} complexes. The globular density element identified in the U2 snRNP region of the complex and surrounded by the HEAT domain of SF3B1 was

assigned to the helicase domain of PRP2 (residues 388–1017). We modelled the helicase in an open state, consistent with it being trapped in a post-translocation state, and identified seven intron nucleotides accommodated by its RNA-binding tunnel. Compared with a previously published yeast B^{act} structure (PDB identifier 7DCO), the helicase domain of PRP2 ($PRP2^{core}$) was no longer positioned on the convex side of $SF3B1^{HEAT}$, but translocated along the intron towards the branch helix. The PRP2-bound RNA formed a continuous density stretch with the intron strand of the U2–BS duplex, showing that the helicase had translocated in a 3′-to-5′ direction from its position in the B^{act} complex. Supporting this model, density for the RES complex subunits RBMX2 and SNIP1, which bind intron regions (or are located close) downstream of the branch helix, could not be observed in B^{AQR}, whereas only a short helical region belonging to the BUD13 subunit (residues 530–557) was identified in the proximity of the MA3 domain of CWC22. Besides modelling the helicase domain of PRP2, we also built de novo a large part of the $PRP2^{NTD}$ that was missing in previous B^{act} structures. The N-terminal most α-helical region (the pin, residues 161–193) was positioned at the periphery of the helicase domain, where it established interfaces with the RecA1, WH and HB domains of the helicase; two other long helices (the clip, residues 223–256 and residues 264–296) were engaged in tight interactions with the PPIL4 linker region and the long helix of SKIP (residues 286–340), respectively. The latter $PRP2^{NTD}$ helix extended further towards PRP8 and finally reached CWC22.

The new density of PRP2 in B^{AQR} coincides with a significant change to the curvature of $SF3B1^{HEAT}$, with its N-terminal HEAT repeats no longer in contact with the BS–U2 helix. However, the density of BS-A was still observed in the SF3B1–PHF5A binding pocket, with the hinge region of SF3B organized as in the closed state. Consistently, the SF3A2 and SF3A3 matrin-type zinc-finger domains were organized and positioned as in the B^{act} complex and still engaged the 5′-end of the U2 snRNA. We modelled SF3B1 by individually docking its consecutive HEAT repeats and manually adjusting their fit to the map. The resulting structure of the reconfigured SF3B1 differed from all other known conformations of $SF3B1^{HEAT}$. Because of the PRP2-induced reconfiguration of $SF3B1^{HEAT}$, its SF3B6 and SRRM1 binding partners were destabilized from their earlier locations on the N-

terminal side of the HEAT superhelix and, compared with B^{act}, their densities were no longer observed in B^{AQR}.

Facing the PRP2 density element and following the RNA density exiting the helicase cassette, a new small globular domain appeared to be recruited to B^{AQR}. We assigned it to the C-terminal RRM region of PPIL4 on the basis of its continuity, through an α-helical density, to the predicted cyclophilin-type PPIase domain that interacts with the long helix of SKIP(286–340) and proteomics analysis. A similar SKIP-bound PPIase density has previously been observed in the later C complex (PDB identifier 5yzg), at an almost identical position; in the published C complex model (PDB 5YZG), the authors assigned it to the PPIase domain of the PPIase PPIG[30]. The latter lacks a RRM domain and is more abundant in C complexes, but only in trace amounts in B^{AQR}. The other PPIase identified in the B^{AQR} proteome, which comprises a PPIase and an RRM domain, is PPIE, an IBC module component located at the periphery of the spliceosome.

In addition to PPIL4 and PPIE, we modelled two other PPIases: PPIL1, interacting with and, probably, stabilizing the PRP19 helical bundle onto the B^{AQR} core and PPIL2. The latter PPIase adopted an extended conformation with its N-terminal tandem U-box motifs, interacting with SNU114 and separated from its C-terminal PPIase domain by about 90 Å. An ordered linker region (residues 234–266) connected the two PPIL2 moieties, as previously observed in a B^{act} cryo-EM structure[5]. We did not observe density for the PPIase domain of CWC27, which is consistent with it being destabilized by the propagated changes induced by the remodelling of SF3B1. However, its interacting partner CWC24 was still present in B^{AQR}, with densities observed for both its zinc finger motif (residues 190–238), sequestering the 5′SS and its C-terminal moiety (residues 262–309) bound to the BPB domain of SF3B3. The final model of the complex (Supplementary Data 2), including its more dynamic peripheral modules, consisted of four RNA molecules (the U2, U5 and U6 snRNA, and the MINX pre-mRNA substrate) and 45 individual polypeptide chains, among which 3 are splicing helicases (PRP2, BRR2 and Aquarius) and 4 are PPIases (PPIL1, PPIL2, PPIL4 and PPIE).

Morphing and generation of movies

The PRP2 helicase translocation trajectory was generated using ChimeraX (v.1.3), with the morph functionality starting from the B^{act} state of the SF3B module (that is, SF3B1 in a closed conformation with PRP2 positioned on the convex side of $SF3B1^{HEAT}$). The NTD domains of SF3B1 ($SF3B1^{NTD}$) and PRP2 ($PRP2^{NTD}$) were not considered in the morphing analysis, and the U2 sequence was limited to the U2 stem-loop IIa structure and the BS-interaction region. The initial (B^{act}) and final (B^{AQR}) trajectory snapshots were aligned using the PHF5A subunit of SF3B as a reference before morphing. The RNA-bound state of PRP2 in B^{act} was modelled based on the yeast B^{act} structure[5]. Supplementary Videos 1–4 were generated using ChimeraX (v.1.3), and Supplementary Video 5 was generated using PyMol.

Helicase assays

The helicase assays were performed as previously described for other splicing helicases[6,56,57,58,59]. To assess the ability of human PRP2 to unwind the U2–BS helix in the presence of its different cofactors, we performed in vitro helicase assays on fluorescently labelled model substrates comprising a perfect double-stranded RNA duplex followed by a single-stranded 3′ overhang. The unwinding activity of PRP2 activity was monitored using either a gel-based readout[6], in which the single-stranded Cy5-labelled product is separated from the helicase substrate on a native polyacrylamide gel, or by recording the time course of decrease in the fluorescence of the substrate[58] as the duplex is unwound. In this case, the dual-labelled RNA strand, displaced because of the helicase activity, forms an intramolecular hairpin that brings in proximity the terminal Cy5 probe and its spectrally overlapping dark quencher (BHQ-2), thereby leading to fluorophore emission quenching and fluorescence decay.

The helicase substrate for the gel-based assay was prepared by mixing 30 μM Cy5-labelled strand (5′-CACCAGCUCCGUAGGCGC-Cy5-3′) with 45 μM unlabelled RNA oligonucleotide (5′-**GCGCCUACGGAGCUGGUG**GCGUAGGCGCAAAAAAAAAAAAAAAAAAA-3′, the complementary region is shown in bold) in 20 mM HEPES-KOH, pH 7.5. The RNA substrate used in the fluorescent-based helicase assay[58] was prepared by mixing in a similar molar ratio the dual-labelled RNA oligonucleotide (5′-Cy5-

GCGCCUACGCCACCAGCUCCGUAGGCGC-BHQ-2-3′) with the unlabelled strand (5′-**GCGCCUACGGAGCUGGUGGCGUAGGCGC**AAAAAAAAAAAAAAAAAAAA-3′, the complementary region is shown in bold) in RNA annealing buffer (6 mM HEPES-KOH, pH 7.5, 50 mM KCl and 0.2 mM $MgCl_2$). In both cases, the single-stranded RNA oligonucleotides were annealed by sequential incubation at 95 °C for 2 min, then at 80 °C for 10 min and then slowly cooled down to room temperature and stored on ice. The RNA and DNA oligonucleotides used in the helicase assays were obtained from IDT (Integrated DNA Technologies) or Microsynth.

In a typical gel-based helicase assay, 50–100 nM fluorescently labelled RNA substrate was mixed in 20 µl on ice with increasing concentrations of PRP2 constructs or PRP2 complexes (1–10 µM) in helicase assay buffer (20 mM HEPES-KOH, pH 7.5, 50 mM KCl, 2 mM $MgCl_2$, 0.1 mg ml^{-1} BSA (NEB), 0.08 U μl^{-1} RNasin (Promega), 5% (v/v) glycerol and 0.5 µM competitor DNA (5′-GCGCCTACGGAGCTGGTG-3′), final concentrations). The GPKOW cofactor was added in a 5-fold molar excess over PRP2(137–1022) or PRP2(137–1022)–PPIL4. After pre-incubation for 10 min at 37 °C (or 20 min at 30 °C), the unwinding reaction was initiated by the addition of 2 mM ATP, and the samples were incubated for 1 h at 37 °C (or 1 h at 30 °C). The reactions were stopped by the addition of 5 µl quenching buffer (0.83 mM Tris-HCl, pH 7.6, 5 mM EDTA, 5% (v/v) glycerol, 0.0025% bromophenol blue, 0.0025% xylene cyanol FF and 0.04 U μl^{-1} proteinase K (NEB), final concentrations) and incubated for 20 min at 37 °C. RNA duplex unwinding was assessed on a 14% polyacrylamide native gel prepared in 1× TBE or 1× Tris-glycine buffer and was run in 0.5× TBE (or 1× Tris-glycine) at 100 V for approximately 90 min at room temperature. The yeast Prp22p helicase, used as a positive control, was prepared as previously described[27]. The RNA gels were scanned at the Cy5 excitation peak using an iBright 1500 imaging system (Invitrogen). The gel-based helicase assays were repeated at least two times.

The molecular beacon helicase assays were performed using a Fluorolog spectrofluorometer (Horiba). In brief, 200 nM dual-labelled substrate was mixed with 1–5 µM purified helicase in reaction buffer (20 mM HEPES-KOH, pH 7.5, 50–150 mM KCl and 2 mM $MgCl_2$) in the absence of ATP

and preincubated for at least 10 min at room temperature. As for the gel-based unwinding assays, GPKOW was added in a 5-fold molar excess over PRP2(137–1022), whereas PPIL4 was added in a 3-fold molar excess. The reactions were subsequently transferred to a 1.5 mm ultra-microcuvette (105.252-QS, Hellma) and equilibrated at the measurement temperature (25 °C or 30 °C) in the spectrometer. The unwinding reaction was initiated by the addition of 2 mM ATP (final concentration), and the decay in Cy5 fluorescence was immediately recorded at an interval time of 1 s and an integration time of 0.1 s. The Cy5 fluorescent probe was excited at 645 nm (3 nm slit width) and its emission was measured at 667 nm (3 nm slit width) and corrected for detector noise. Data were analysed and plotted using OriginPro 2020 (9.7.0.188). The fluorescence-based helicase assay was repeated at least three times. Three independent preparations of PRP2(137–1022) were tested in these assays.

Electrophoretic mobility shift assay

To investigate the RNA binding activity of the different PRP2 complexes tested in the unwinding assays, 50 nM fluorescent substrate was mixed with increasing concentrations of helicase or helicase cofactor samples (0.1–8 μM). The 20 μl binding reactions were prepared in RNA binding buffer (20 mM HEPES-KOH, pH 7.5, 50 mM KCl, 2 mM $MgCl_2$, 0.1 mg ml^{-1} BSA (NEB) and 5% (v/v) glycerol, final concentrations) and incubated for 30 min at room temperature before being loaded on a 5% polyacrylamide native gel prepared in 1× Tris-glycine buffer. The native RNA gel was pre-run at 40 V for 30 min at room temperature and run then at 60 V in 1× Tris-glycine. The gels were scanned at the Cy5 excitation peak. The RNA binding assays were repeated at least two times, and similar results were obtained.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The coordinate files have been deposited into the PDB with the identifiers 7QTT (high-resolution core) and 8CH6 (overall composite model). The cryo-EM maps have been deposited into the Electron Microscopy Data Bank with the identifiers EMD-14146 (high-resolution core) and EMD-16658 (overall reconstruction). The mass spectrometry protein composition and model building of the B^{AQR} spliceosomes are available at FigShare (https://doi.org/10.6084/m9.figshare.22047275).

References

1. Warkocki, Z. et al. Reconstitution of both steps of *Saccharomyces cerevisiae* splicing with purified spliceosomal components. *Nat. Struct. Mol. Biol.* **16**, 1237–1243 (2009).

2. Zhang, X. et al. Structure of the human activated spliceosome in three conformational states. *Cell Res.* **28**, 307–322 (2018).

3. Rauhut, R. et al. Molecular architecture of the *Saccharomyces cerevisiae* activated spliceosome. *Science* **353**, 1399–1405 (2016).

4. De Bortoli, F., Espinosa, S. & Zhao, R. DEAH-box RNA helicases in pre-mRNA splicing. *Trends Biochem. Sci.* **46**, 225–238 (2021).

5. Bai, R. et al. Mechanism of spliceosome remodeling by the ATPase/helicase Prp2 and its coactivator Spp2. *Science* https://doi.org/10.1126/science.abe8863 (2021).

6. De, I. et al. The RNA helicase Aquarius exhibits structural adaptations mediating its recruitment to spliceosomes. *Nat. Struct. Mol. Biol.* **22**, 138–144 (2015).

7. De, I., Schmitzova, J. & Pena, V. The organization and contribution of helicases to RNA splicing. *Wiley Interdiscip. Rev. RNA* **7**, 259–274 (2016).

8. Kastner, B., Will, C. L., Stark, H. & Luhrmann, R. Structural insights into nuclear pre-mRNA splicing in higher eukaryotes. *Cold Spring*

Harb. Perspect. Biol. https://doi.org/10.1101/cshperspect.a032417 (2019).

9. Wan, R., Bai, R., Zhan, X. & Shi, Y. How is precursor messenger RNA spliced by the spliceosome? *Annu. Rev. Biochem.* **89**, 333–358 (2020).

10. Wilkinson, M. E., Charenton, C. & Nagai, K. RNA splicing by the spliceosome. *Annu. Rev. Biochem.* **89**, 359–388 (2020).

11. Galej, W. P. et al. Cryo-EM structure of the spliceosome immediately after branching. *Nature* **537**, 197–201 (2016).

12. Wan, R., Yan, C., Bai, R., Huang, G. & Shi, Y. Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science* **353**, 895–904 (2016).

13. Haselbach, D. et al. Structure and conformational dynamics of the human spliceosomal B^{act} complex. *Cell* **172**, 454–464.e11 (2018).

14. King, D. S. & Beggs, J. D. Interactions of PRP2 protein with pre-mRNA splicing complexes in *Saccharomyces cerevisiae*. *Nucleic Acids Res.* **18**, 6559–6564 (1990).

15. Bessonov, S. et al. Characterization of purified human B^{act} spliceosomal complexes reveals compositional and morphological changes during spliceosome activation and first step catalysis. *RNA* **16**, 2384–2403 (2010).

16. Bertram, K. et al. Structural insights into the roles of metazoan-specific splicing factors in the human step 1 spliceosome. *Mol. Cell* **80**, 127–139.e6 (2020).

17. Wan, R., Bai, R., Yan, C., Lei, J. & Shi, Y. Structures of the catalytically activated yeast spliceosome reveal the mechanism of branching. *Cell* **177**, 339–351.e13 (2019).

18. Schmitt, A., Hamann, F., Neumann, P. & Ficner, R. Crystal structure of the spliceosomal DEAH-box ATPase Prp2. *Acta Crystallogr. D Struct. Biol.* **74**, 643–654 (2018).

19. Liu, H. L. & Cheng, S. C. The interaction of Prp2 with a defined region of the intron is required for the first splicing reaction. *Mol. Cell. Biol.* **32**, 5056–5066 (2012).

20. Dziembowski, A. et al. Proteomic analysis identifies a new complex required for nuclear pre-mRNA retention and splicing. *EMBO J.* **23**, 4847–4856 (2004).

21. Bao, P., Will, C. L., Urlaub, H., Boon, K. L. & Luhrmann, R. The RES complex is required for efficient transformation of the precatalytic B spliceosome into an activated B^{act} complex. *Genes Dev.* **31**, 2416–2429 (2017).

22. Semlow, D. R., Blanco, M. R., Walter, N. G. & Staley, J. P. Spliceosomal DEAH-box ATPases remodel pre-mRNA to activate alternative splice sites. *Cell* **164**, 985–998 (2016).

23. Bao, P., Hobartner, C., Hartmuth, K. & Luhrmann, R. Yeast Prp2 liberates the 5′ splice site and the branch site adenosine for catalysis of pre-mRNA splicing. *RNA* **23**, 1770–1779 (2017).

24. Sloan, K. E. & Bohnsack, M. T. Unravelling the mechanisms of RNA helicase regulation. *Trends Biochem. Sci.* **43**, 237–250 (2018).

25. Bono, F., Ebert, J., Lorentzen, E. & Conti, E. The crystal structure of the exon junction complex reveals how it maintains a stable grip on mRNA. *Cell* **126**, 713–725 (2006).

26. Kim, S. H., Smith, J., Claude, A. & Lin, R. J. The purified yeast pre-mRNA splicing factor PRP2 is an RNA-dependent NTPase. *EMBO J.* **11**, 2319–2326 (1992).

27. Warkocki, Z. et al. The G-patch protein Spp2 couples the spliceosome-stimulated ATPase activity of the DEAH-box protein Prp2 to catalytic activation of the spliceosome. *Genes Dev.* **29**, 94–107 (2015).

28. Schwer, B. A conformational rearrangement in the spliceosome sets the stage for Prp22-dependent mRNA release. *Mol. Cell* **30**, 743–754 (2008).

29. Chung, C. S. et al. Dynamic protein–RNA interactions in mediating splicing catalysis. *Nucleic Acids Res.* **47**, 899–910 (2019).

30. Zhan, X., Yan, C., Zhang, X., Lei, J. & Shi, Y. Structure of a human catalytic step I spliceosome. *Science* **359**, 537–545 (2018).

31. Zhan, X., Lu, Y., Zhang, X., Yan, C. & Shi, Y. Mechanism of exon ligation by human spliceosome. *Mol. Cell* **82**, 2769–2778.e4 (2022).

32. Bertram, K. et al. Cryo-EM structure of a pre-catalytic human spliceosome primed for activation. *Cell* **170**, 701–713.e11 (2017).

33. Fica, S. M., Oubridge, C., Wilkinson, M. E., Newman, A. J. & Nagai, K. A human postcatalytic spliceosome structure reveals essential roles of metazoan factors for exon ligation. *Science* **363**, 710–714 (2019).

34. Gao, K., Masuda, A., Matsuura, T. & Ohno, K. Human branch point consensus sequence is yUnAy. *Nucleic Acids Res.* **36**, 2257–2267 (2008).

35. Taggart, A. J. et al. Large-scale analysis of branchpoint usage across species and cell lines. *Genome Res.* **27**, 639–649 (2017).

36. Berget, S. M. Exon recognition in vertebrate splicing. *J. Biol. Chem.* **270**, 2411–2414 (1995).

37. Yeo, G., Hoon, S., Venkatesh, B. & Burge, C. B. Variation in sequence and organization of splicing regulatory elements in vertebrate genes. *Proc. Natl Acad. Sci. USA* **101**, 15700–15705 (2004).

38. Hirose, T. et al. A spliceosomal intron binding protein, IBP160, links position-dependent assembly of intron-encoded box C/D snoRNP to pre-mRNA splicing. *Mol. Cell* **23**, 673–684 (2006).

39. Yan, C. et al. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science* **349**, 1182–1191 (2015).

40. Kim, S. H. & Lin, R. J. Spliceosome activation by PRP2 ATPase prior to the first transesterification reaction of pre-mRNA splicing. *Mol. Cell.*

Biol. **16**, 6810–6819 (1996).

41. Ohrt, T. et al. Prp2-mediated protein rearrangements at the catalytic core of the spliceosome as revealed by dcFCCS. *RNA* **18**, 1244–1256 (2012).

42. Kaur, H., Groubert, B., Paulson, J. C., McMillan, S. & Hoskins, A. A. Impact of cancer-associated mutations in Hsh155/SF3b1 HEAT repeats 9–12 on pre-mRNA splicing in *Saccharomyces cerevisiae*. *PLoS ONE* **15**, e0229315 (2020).

43. Carrocci, T. J., Zoerner, D. M., Paulson, J. C. & Hoskins, A. A. *SF3b1* mutations associated with myelodysplastic syndromes alter the fidelity of branchsite selection in yeast. *Nucleic Acids Res.* https://doi.org/10.1093/nar/gkw1349 (2017).

44. Yeh, T.-C. et al. Splicing factor Cwc22 is required for the function of Prp2 and for the spliceosome to escape from a futile pathway. *Mol. Cell. Biol.* **31**, 43–53 (2011).

45. Edwalds-Gilbert, G., Kim, D. H., Silverman, E. & Lin, R. J. Definition of a spliceosome interaction domain in yeast Prp2 ATPase. *RNA* **10**, 210–220 (2004).

46. Cretu, C. et al. Molecular architecture of SF3b and structural consequences of its cancer-related mutations. *Mol. Cell* **64**, 307–319 (2016).

47. Cretu, C. et al. Structural basis of intron selection by U2 snRNP in the presence of covalent inhibitors. *Nat. Commun.* **12**, 4491 (2021).

48. Tholen, J., Razew, M., Weis, F. & Galej, W. P. Structural basis of branch site recognition by the human spliceosome. *Science* **375**, 50–57 (2022).

49. Will, C. L. et al. Characterization of novel SF3b and 17S U2 snRNP proteins, including a human Prp5p homologue and an SF3b DEAD-box protein. *EMBO J.* **21**, 4978–4988 (2002).

50. Bieniossek, C., Imasaki, T., Takagi, Y. & Berger, I. MultiBac: expanding the research toolbox for multiprotein complexes. *Trends Biochem. Sci.* **37**, 49–57 (2012).

51. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat. Methods* **16**, 1146–1152 (2019).

52. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).

53. Jakobi, A. J., Wilmanns, M. & Sachse, C. Model-based local density sharpening of cryo-EM maps. *eLife* **6**, e27131 (2017).

54. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).

55. Terwilliger, T. C., Sobolev, O. V., Afonine, P. V. & Adams, P. D. Automated map sharpening by maximization of detail and connectivity. *Acta Crystallogr. D Struct. Biol.* **74**, 545–559 (2018).

56. Studer, M. K., Ivanovic, L., Weber, M. E., Marti, S. & Jonas, S. Structural basis for DEAH-helicase activation by G-patch proteins. *Proc. Natl Acad. Sci. USA* **117**, 7159–7170 (2020).

57. Christian, H., Hofele, R. V., Urlaub, H. & Ficner, R. Insights into the activation of the helicase Prp43 by biochemical studies and structural mass spectrometry. *Nucleic Acids Res.* **42**, 1162–1179 (2014).

58. Hamann, F. et al. The structure of Prp2 bound to RNA and ADP-BeF_3^- reveals structural features important for RNA unwinding by DEAH-box ATPases. *Acta Crystallogr. D Struct. Biol.* **77**, 496–509 (2021).

59. Felisberto-Rodrigues, C. et al. Structural and functional characterisation of human RNA helicase DHX8 provides insights into the mechanism of RNA-stimulated ADP release. *Biochem. J.* **476**, 2521–2543 (2019).

60. Zhang, Z. et al. Molecular architecture of the human 17S U2 snRNP. *Nature* **583**, 310–313 (2020)

61. Yan, C., Wan, R. & Shi, Y. Molecular mechanisms of pre-mRNA splicing through structural biology of the spliceosome. *Cold Spring Harb. Perspect. Biol.* **11**, a032409 (2019).

62. Cretu, C. et al. Structural basis of splicing modulation by antitumor macrolide compounds. *Mol. Cell* **70**, 265–273.e8 (2018).

63. Beusch, I. et al. Targeted high throughput mutagenesis of the human spliceosome reveals its in vivo operating principles. Preprint at *bioRxiv* https://doi.org/10.1101/2022.11.13.516350 (2022).

64. Zhang, J. et al. DHX15 is involved in SUGP1-mediated RNA missplicing by mutant SF3B1 in cancer. *Proc. Natl Acad. Sci. USA* **119**, e2216712119 (2022).

65. Feng, Q. et al. Splicing quality control mediated by DHX15 and its G-patch activator, SUGP1. Preprint at *bioRxiv* https://doi.org/10.1101/2022.11.14.516533 (2022).

Acknowledgements

We thank C. Richardson (The Institute of Cancer Research, London) for technical support in computation; M. Raabe, S. König and A. Chernev (Max-Planck Institute, Göttingen) for mass spectrometry analysis; S. Bessonov (Max-Planck Institute, Göttingen) for technical advice in splicing biochemistry; and A. Hoskins (University of Wisconsin-Madison, USA) for critical reading of the manuscript. This work was supported by the German Research Foundation (grant DFG PE 2079/2-2 and DFG PE 2079/4-1), the Wellcome Trust (220300Z/20/Z) and the Institute of Cancer Research.

Author information

Author notes

1. Jana Schmitzová

Present address: Molecular Biology, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany

2. Constantin Cretu

   Present address: Cluster of Excellence Multiscale Bioimaging (MBExC), Universitätsmedizin Göttingen, Göttingen, Germany

3. These authors contributed equally: Jana Schmitzová, Constantin Cretu

Authors and Affiliations

1. Macromolecular Crystallography, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany

   Jana Schmitzová, Constantin Cretu & Vladimir Pena

2. Research Group Mechanisms and Regulation of Splicing, The Institute of Cancer Research, London, UK

   Constantin Cretu & Vladimir Pena

3. Molecular Biology, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany

   Christian Dienemann

4. Bioanalytical Mass Spectrometry, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany

   Henning Urlaub

5. Institute of Clinical Chemistry, Bioanalytics, University Medical Center Sciences, Göttingen, Germany

   Henning Urlaub

Contributions

J.S. established purification of the spliceosomes, prepared grids and collected cryo-EM data, processed cryo-EM data, reconstructed initial maps and built initial models by docking. C.C. performed advanced processing of cryo-EM data, which resulted in the final deposited maps. C.C. built and refined the structural models. J.S. and C.C. purified recombinant proteins. C.C. cloned the expression constructs, performed the interaction assays and helicase assays with recombinant proteins. C.D. assisted with cryo-EM data collection and early data processing. H.U. performed mass spectrometry analyses. C.C. and V.P. wrote the paper, with input from J.S. and C.D. V.P. supervised the research.

Corresponding author

Correspondence to Vladimir Pena.

Ethics declarations

Competing interests

The authors declare no competing interests.

Peer review

Peer review information

Nature thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

Additional information

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Extended data figures and tables

Extended Data Fig. 1 Purification, biochemical analysis and cryo-EM image-processing workflow of B^{AQR}.

a, Schematic depicting the MINX pre-mRNA substrate, fused to three downstream MS2 aptamers, which was used to assemble the B^{AQR} spliceosome in the HeLa nuclear extract. **b**, Pre-mRNA splicing is arrested by incorporating the recombinant dominant-negative IBC/Aquarius^{K829A} in the human spliceosomes. Typical *in vitro* splicing reactions of the model MINX pre-mRNA substrate were set up in the presence (+IBC^{K829A}) or absence (+buffer) of the recombinant complex added *in trans*, in excess over the endogenous IBC. The reactions were stopped after 0–120 min and the Cy5-labeled RNA substrate was extracted, analyzed on a denaturing polyacrylamide gel, and visualized by in-gel fluorescence. The experiment was repeated at least three times, with sample resulting from distinct preparations, with similar results. **c**, Protein and RNA composition of the purified B^{AQR} complexes. The SDS-PAGE gels were stained with Coomassie (left) and SYBR Gold (right). The purification was repeated at least 3 times, with similar results. For gel source data, see Supplementary Fig. 1. **d**, Representative cryo-EM micrograph of the stalled B^{AQR} spliceosome and typical reference-free 2D class averages of B^{AQR} particle images. We collected two independent datasets from two different sample preparations **e**, Cryo-EM data processing schematic showing the most significant steps of the computational analysis. **f–i**, Global resolution of key B^{AQR} maps (see also Extended Data Fig. 2). The FSC (Fourier Shell Correlation) analysis was carried out in RELION 3.1 (left). The angular distribution (right) of the B^{AQR} particles contributing to the map shown on top is depicted in two different orientations, with the red color indicating a higher number of particles at a given projection angle.

Extended Data Fig. 2 Cryo-EM image-processing by local classification of the peripheral modules of the B^{AQR} complex.

Cryo-EM data processing schematic showing the most significant steps of the computational analysis, the FSC (Fourier Shell Correlation) analysis of

the obtained focused maps, and the angular distribution of the particles contributing to the final volumes.

Extended Data Fig. 3 Cryo-EM maps and structure of the B^{AQR} spliceosome.

a, Overall maps of state A and B complexes depicted together with the final model. **b**, Local resolution of the state A maps, estimated in RELION and visualized in ChimeraX. **c**, Cryo-EM density snapshots. Various subunits are colored and labeled. **d**, Selected snapshots of modeled B^{AQR} subunits are depicted together with the unsharpened map of the core region of the complex (map M2). **e**, Structural comparison between B^{act}, B^{AQR}, and C complexes. PRP2, Aquarius, SYF1, the U2 snRNA, and the pre-mRNA substrate are colored and labeled. Note the large-scale repositioning of the PRP2 RNA helicase during the conversion of B^{act} to B^{AQR}, and of the helicase Aquarius during the transition of B^{act}/B^{AQR} complexes to the C-stage spliceosome.

Extended Data Fig. 4 Functionally-important contacts in the B^{AQR} spliceosome.

a–f, Interactions between PRP2, PRP8, SF3B1, CWC22 and SKIP. Interacting distances corresponding to polar and hydrophobic contacts are indicated as dashed lines. **g**, Overview of PRP2 and the composite molecular brake. PRP2's conserved domains are color-coded. HB – helix-bundle; OB – oligonucleotide/oligosaccharide- binding fold; WH – winged-helix. **h**, Interactions between PPIL4, SKIP, and PRP2. Residues involved in contacts are depicted as spheres and shown in the same color as their interacting partner. **i**, Interactions between the wedge element of SKIP and $PRP2^{core}$. **j**, The wedge element of SKIP appears to lock the RecA domains of PRP2. The ADP molecule is modeled by superposition with *Ct* PRP2 (PDB 6zm2) and is shown for the sake of orientation. **k**, Superposition between human PRP2 in the open conformation (observed in B^{AQR}) and *Ct* PRP2 in the closed conformation (colored in teal) (PDB 6zm2). Equivalent residues of PRP2 interacting with SKIP in B^{AQR} are shown for *Ct* PRP2. Structures from panels i and k are depicted in the same orientation. The red arrow

indicates the movement of the RecA2-like domain during the helicase transition from the open to the closed conformation. **l**, Genetic interactions from yeast mapped on the structure of human B^{AQR}. The cold-sensitive allele of $Prp2p^{Q548N}$ genetically interacts with the D450G and V502F substitutions of Hsh155p. Residues involved in genetic interactions in yeast (blue) are depicted as spheres and mapped on the B^{AQR} model. **m-o**, Structural superposition of the human (*Hs*) and budding yeast (*Sc*) MA3 domain of CWC22, composed from HEAT repeats. The 10th HEAT repeat (residues 454–491) of Cwc22p is required for the productive function of yeast Prp2p and interacts with the "hook" motif of $PRP2^{NTD}$ and BUD13 in B^{AQR}. This suggests that Prp2p, Cwc22p, and Bud13p may interact in a similar fashion in budding yeast spliceosomes, thus explaining the functional connection between Prp2p and Cwc22p. Note that human CWC22 residues that interact with PRP2 and BUD13 are conserved in budding yeast Cwc22p (shown in n and o). The protein subunits, U2 snRNA and the pre-mRNA substrate are colored and labeled.

Extended Data Fig. 5 Estimation of the number of nucleotides translocated by PRP2 during the B^{act}-to-B^{AQR} transition.

a,b, Superposition between budding yeast (PDB 7DCO) and human (PDB 5Z56) B^{act} complexes over equivalent residues from PHF5A. For clarity's sake, only SF3B1, PRP2, the intron and U2 snRNA are depicted. All subunits are labeled accordingly. **c**, PRP2:RNA from human B^{AQR}, as the only available structure of the human counterpart, was superimposed onto the Prp2p helicase from budding yeast B^{act} and shown in the same orientation as in b. Note that the conformation of the RNA strand bound by human PRP2 or budding yeast Prp2p is virtually unchanged. **d**, Assignment of nucleotides bound by human PRP2 in B^{act} based on the superposition. **e**, PRP2 translocates about 19 nucleotides towards the branch helix during the B^{act}-to-B^{AQR} transition. PRP2, RBMX2 and PPIL4 positions on the intron in different spliceosomal complexes are depicted. **f**, PRP2 translocation results in a substantial change in the intron's conformation, dissociation of RBMX2 and recruitment of PPIL4 to the intron.

Extended Data Fig. 6 Conformational transitions of SF3B1.

a, Destabilization of SF3B6, de-structuring of SF3B1^{NTD}, and the reorientation of PRP8's EN (endonuclease-like, PRP8^{EN}) and RH (RNase H-like, PRP8^{RH}) domains upon PRP2 translocation. PRP2 is located on the plane above SF3B1^{HEAT} and is not shown for clarity's sake. The reactive BS-A, the guanosine of the 5'SS, and the catalytic metal ions are shown as spheres and labeled. **b**, Close-up view of SF3B1, the intron, and the U2 snRNA from B^{AQR} (color-coded) and B^{act} (grey and black) after superposition of SF3B1's equivalent residues. **c**, The binding pocket of the BS-A (primary pocket) is virtually unchanged in B^{act} and B^{AQR}. **d**, Cycle of SF3B1 transitions in splicing. The conformational transitions of SF3B1 are depicted based on cryo-EM structures and biochemical analyses. Here we refer to the intermediates II and III as pre-A1 and pre-A2 for clarity. The key spliceosome complexes representative for the SF3B1 intermediates shown here are: 17S U2 snRNP[48,60], pre-A1 (A-like cross-exon complex bound by spliceostatin A[47]), pre-A2 (ref. [48]), A-to-B^{act} (reviewed in refs. [8,10,61]), B^{AQR} (this work), the SF3B complex[46,62]. Helicases that facilitate the transitions are shown in red. The helicase DHX15, colored magenta, mediates the disassembly of kinetically-slowed complexes (e.g. formed on suboptimal introns, weak splice sites and PPTs, multiple branch sites or cryptic sites[63,64,65]).

Extended Data Fig. 7 Biochemical analysis of human PRP2.

a, Size-exclusion chromatography (SEC) profiles of PRP2 (137-1022), PPIL4 and the mixture of the two proteins shows that PRP2 (137-1022) forms a stable complex with PPIL4, in an RNA-independent manner. PRP2 (137-1022) comprises both the modeled PRP2^{NTD} (*i.e.*, the pin, clip, and hook elements) and the helicase core. **b**, SDS-PAGE gels corresponding to the fractions from **a**. The *in vitro* reconstitution experiments were repeated two times, using two independent preparations of PRP2 (137–1022) and PPIL4. **c**, SDS-PAGE gel of SEC fractions showing the reconstitution of a stable PRP2-PPIL4 complex *in vivo* by co-expression in insect cells. PRP2 (137–1022) and PPIL4 were co-expressed in Sf9 insect cells from individual baculoviruses. The complex was captured by Strep-Tactin affinity followed by SEC. The preparation was performed twice with similar results **d**, The RNA-binding activity of human PRP2 in the presence of GPKOW and

PPIL4. The ability of PRP2 (137–1022) and of the purified PRP2 (137–1022)-PPIL4 and PRP2 (137-1022)-GPKOW complexes to bind a Cyanine 5 (Cy5)-labeled RNA substrate was evidenced by EMSA. The RNA substrate comprised an RNA duplex, followed by 30 nucleotides 3′ single-stranded overhang, mimicking PRP2's spliceosome substrate observed in B^{AQR}. The free RNA substrate was separated from PRP2-bound (or cofactor-bound) species on a 5% polyacrylamide native gel. The EMSA gels were imaged at the Cy5 excitation peak. The assays were repeated two times. **e**, The RNA binding activity of human PPIL4 in the absence and presence of PRP2 (137–1022). Compared to the EMSA shown in **d**, PPIL4 was added in a 5-fold excess over PRP2 (137–1022) and PRP2's final assay concentrations are indicated above the last three gels lanes. Note the apparent increase in PRP2's RNA affinity in the presence of PPIL4. The EMSAs were repeated three times. **f**, SEC profiles showing the formation of a stable complex between PRP2 (137–1022) and GPKOW, in an RNA-independent manner. The SDS-PAGE corresponding to the complex purified by SEC is shown. **g**, The RNA-binding activity of human GPKOW assessed by an EMSA. The same RNA substrate was used as in **d** and the assay was repeated four times. **h**, The helicase activity of human PRP2 was investigated using a fluorescence-based assay. Compared to the gel-based helicase assays shown in **i-l**, the fluorescence-based assay employs a dual-labeled helicase substrate. Displacement of the labeled strand by the helicase leads to the formation of an intramolecular hairpin, which brings in proximity the Cy5 fluorophore and its spectrally overlapping quencher (BHQ-2). The decrease in substrate's fluorescence upon unwinding is monitored as a function of time. Several representative fluorescence traces of PRP2, recorded under different experimental conditions and in the presence/absence of helicase cofactors, are shown together with the unwinding curves of Prp22p, used as a positive control. **i**, The helicase/unwinding activity of human PRP2 and of the PRP2-GPKOW complex. To assess the ability of PRP2 (137–1022) and of the purified, *in vitro* reconstituted PRP2 (137–1022)-GPKOW complex to unwind RNA-RNA duplexes, the purified protein samples were mixed with the Cy5-labeled helicase substrate (depicted on the right with the labeled strand colored in red) in the presence (or absence) of ATP and of a competitor DNA (green). Following a 1-hour incubation, the samples were analyzed on a native 14% polyacrylamide gel and imaged by in-gel fluorescence. Budding yeast Prp22p was used as a positive control. All gel-

based helicase assays were repeated at least two times. **j**, The helicase activity of the *in vivo* reconstituted PRP2 (137–1022)-PPIL4 complex, in the absence or presence of GPKOW. **k**, The helicase activity of human PRP2 (137–1022) in the presence or absence of its cofactor GPKOW. Compared to the assay shown in **i**, the GPKOW cofactor was added to the purified helicase in a 5-fold excess. The concentrations indicated above the native gel represent the final assay concentrations of PRP2 (137–1022) or Prp22p. The RNA bands labeled with an asterisk represent, most likely, degradation products. **l**, Comparative helicase activities of the PRP2 (137–1022)-PPIL4 and PRP2 (137–1022)-GPKOW complexes. For gel source data, see Supplementary Figs. 2 and 3.

Extended Data Fig. 8 Remodeling events at the transition from B^{act} to B^{AQR} and then to C complexes.

a, During the B^{act} (PDB 5Z57) to the B^{AQR} transition, translocation of PRP2 in a 3′-to-5′ direction results in the destabilization of the RES complex (RBMX2, BUD13, SNIP1). In addition, SRRM1 and SF3B6/p14 are no longer observed in B^{AQR} due to PRP2-induced SF3B1 opening. During the remodeling of B^{AQR} and transition to the C complex (PDB 5yzg, PDB 6zym), the remaining SF3B/SF3A subunits and CWC24 are released from the branch helix and the 5’SS. The branch helix then moves to the catalytic center, bringing the BS-A and the 5’SS GU nucleotides in proximity for the branching reaction. The different human spliceosome states were structurally aligned by using the PRP8 subunit as a reference and are depicted in two different orientations. The spliceosome subunits are color-coded and shown in cartoon representation. Spliceosome subunits not undergoing significant rearrangement were omitted for the sake of simplicity. **b**, Repositioning of the branch helix (U2/BS) during the transition from B^{act} to B^{AQR} and then to C complexes. The different spliceosome states, B^{act} (PDB 5Z57), B^{AQR} (this work), and C complex (PDB 5yzg, PDB 6zym), were aligned using the PRP8 subunit as a reference. All protein subunits, except PRP8, were omitted and the RNA moieties were color-coded. The reactive BS-A and the 5’SS GU nucleotides are shown as spheres and colored red and light green, respectively.

Extended Data Fig. 9 Aquarius promotes the transfer of the BS-A to the catalytic center during B^{AQR}-to-C complex transition, likely by inducing the "loose" to "open" conformational transition of SF3B1.

a, Lines of structural communication between Aquarius and the branch duplex in B^{AQR}. Aquarius and PRP2 are in diametrically-opposed locations of the B^{AQR} spliceosome. A continuous bridge of proteins is present between Aquarius and the SF3B1:branch duplex. Most of these proteins are SF3A/B subunits and PPIE. Another bridge that primarily involves the RBM22 protein is between Aquarius and U6 catalytic core. The intron region not visible in the density is dashed. Aquarius is depicted in red (RecA-like domains) and light blue (accessory domains). The first nucleotide of the intron (G^{+1}) is positioned in the catalytic center. Other subunits of the spliceosome, including SYF1 and ISY1, are not shown for the sake of clarity **b**, Subunits of the C complex (pdb 5yzg) are shown in the same orientation as in **a**. The U6 snRNA was used as a reference for the superposition between the C and B^{AQR} complexes. Note that PRP2 and the SF3A/B complex have dissociated and the BS-A has been relocated to the catalytic centre (see also **c**, below). RBM22's orientation has remained virtually the same in B^{AQR} and C complexes. The subunits are colored as in a and labeled. The B^{AQR} and C complexes were superimposed over PRP8 (not shown) and U6 snRNA **c**, The BS-A juxtaposed to the first nucleotide of the intron in the catalytic center of the C complex. The catalytic metal ions are shown as magenta spheres. **d**, Superposition between the "open" (observed in the apo SF3B complex, PDB 5IFE) and the "loose" states (observed in B^{AQR}) of SF3B1. Except for $SF3B1^{OPEN}$, all shown subunits belong to the B^{AQR} complex. **e**, B^{AQR} and the SF3B complex in its apo form (PDB 5IFE) were superimposed over equivalent residues from PHF5A (not shown in the figures d, f and g for clarity's sake). Note that the HEAT repeats H16-H20 of SF3B1 adopt virtually identical conformations in the "loose" and "open" states, while the other helical repeats are substantially reorganized. **f–g**, The BS-A's release from the primary pocket of SF3B1 likely induces rearrangement of the HEAT repeats upon the "loose" to "open" conformational transition. Consequently, the contacts between HEAT repeats and PRP8, and those between SF3B2 (attached to the HEAT repeats)

and the U6 snRNA might get disrupted, causing the complete dissociation of the SF3A/B complexes from the spliceosome.

Extended Data Table 1 Data collection, model refinement and validation

Supplementary information

Supplementary Information

Supplementary figures and legends for the supplementary figures, the supplementary data files and the supplementary video files.

Reporting Summary

Peer Review File

Supplementary Data 1

Supplementary Data 2

Supplementary Video 1

Supplementary Video 2

Supplementary Video 3

Supplementary Video 4

Supplementary Video 5

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give

appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

Reprints and Permissions

About this article

Cite this article

Schmitzová, J., Cretu, C., Dienemann, C. *et al.* Structural basis of catalytic activation in human splicing. *Nature* **617**, 842–850 (2023). https://doi.org/10.1038/s41586-023-06049-w

- Received: 29 May 2022
- Accepted: 04 April 2023
- Published: 10 May 2023
- Issue Date: 25 May 2023
- DOI: https://doi.org/10.1038/s41586-023-06049-w

Share this article

Anyone you share the following link with will be able to read this content:

Get shareable link

Sorry, a shareable link is not currently available for this article.

Copy to clipboard

Provided by the Springer Nature SharedIt content-sharing initiative

This article is cited by

- **A clue to the catalytic activation of splicing**
    - Soo-Chen Cheng

  Nature (2023)

A clue to the catalytic activation of splicing

- Soo-Chen Cheng

Nature News & Views 10 May 2023

This article was downloaded by **calibre** from https://www.nature.com/articles/s41586-023-06049-w

Collections

- **Quantum computing** [24 May 2023]
  • From smartphones to supercomputers, modern computers work by manipulating digital 1s and 0s.
- **Spinoff Prize 2023** [24 May 2023]
  • Academic research constantly creates and disperses seeds of scientific insight.



Quantum computers: what are they good for?

For now, absolutely nothing. But researchers and firms are optimistic about the applications.

- Michael Brooks

Spotlight24 May 2023 Nature



Fast tracking quantum-computing tech

Quantum computers are fast approaching the point where they will start benefiting society — and researchers in Japan want to get there as quickly as possible.

Advertisement Feature24 May 2023

This article was downloaded by **calibre** from https://www.nature.com/collections/ghdfdeigcd



The Spinoff Prize 2023

Nature introduces the finalists and those that made it on to the longlist for this year's award.

- Herb Brody

Outlook24 May 2023 Nature



The Spinoff Prize: where are they now?

Clinical trials, industry partnerships and other milestones reached by previous finalists of The Spinoff Prize.

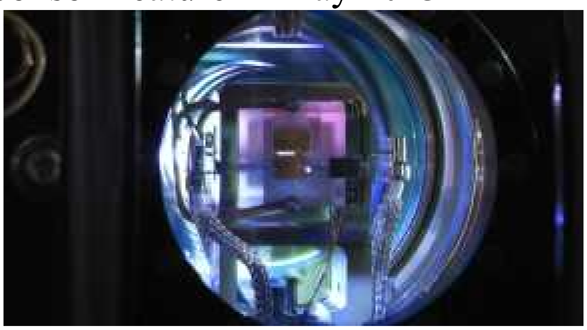
- Neil Savage

Outlook24 May 2023 Nature



[Merck Science Prizes, Awards, Grants and Challenges]

Sponsor Feature24 May 2023



[Commercializing quantum computers step by step]

Alongside developing a quantum computer, one group of scientists is selling its components to other researchers.

- Edwin Cartlidge

Outlook24 May 2023 Nature



A cleaner route to ammonia

A method driven by renewable energy could end the need for fossil fuels in fertilizer production.

- James Mitchell Crow

Outlook24 May 2023 Nature

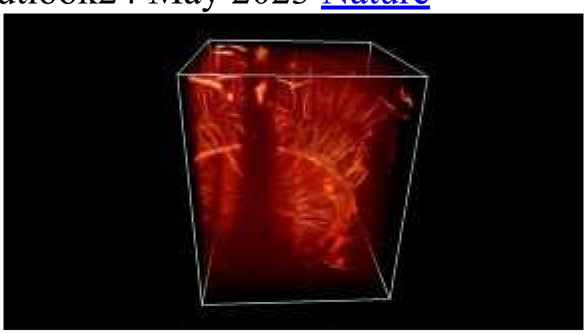


Rewriting the quantum-computer blueprint

An architecture for quantum computers based on parity is attracting money from government and industry.

- Edwin Cartlidge

Outlook24 May 2023 Nature



Revealing vascular roadblocks in the brain

High-resolution imaging quickly identifies blood clots before they inflict major damage.

- Michael Eisenstein

Outlook24 May 2023 Nature



Bioglue breakthrough

A nature-inspired adhesive offers hope for wound healing and haemorrhage control.

- Elie Dolgin

Outlook24 May 2023 Nature

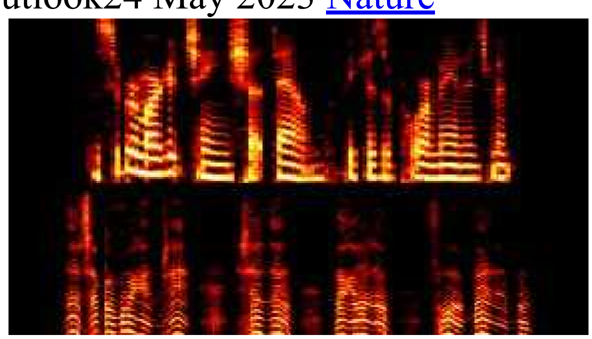


Better batteries built using existing technology

Advances in materials yield safer, cheaper and denser energy storage.

- Neil Savage

Outlook24 May 2023 Nature



Listening for neurological symptoms

Unusual vocal patterns can give clues that help to detect conditions such as Alzheimer’s disease.

- Michael Eisenstein

Outlook24 May 2023 Nature



Accelerating drug development with 3D neural models

A reproducible cell-culture system could help to evaluate new therapies for disorders that affect the brain.

- Christine Evans-Pughe

Outlook24 May 2023 Nature



How to lower carbon levels using light

Photocatalysis using light-emitting diodes could reduce the amount of carbon dioxide emitted by industrial chemical processes.

- Neil Savage

Outlook24 May 2023 Nature

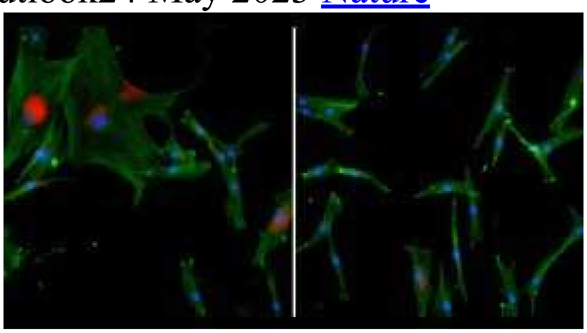


A fat-blocking drug could help to fight metastatic cancer

The antibody drug is moving closer to being tested in people with advanced stages of cancer.

- Elie Dolgin

Outlook24 May 2023 Nature



RNA splicing targets age-related diseases

Manipulating genetic molecules could return cells to a younger state.

- Christine Evans-Pughe

Outlook24 May 2023 Nature

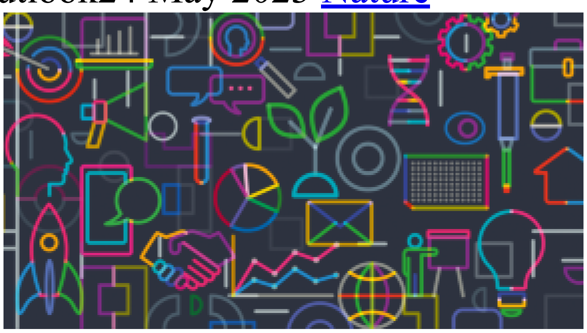


[Disrupting protein folding to tackle cancer]

Intermediate states show promise as drug targets.

- Elie Dolgin

Outlook24 May 2023 [Nature]

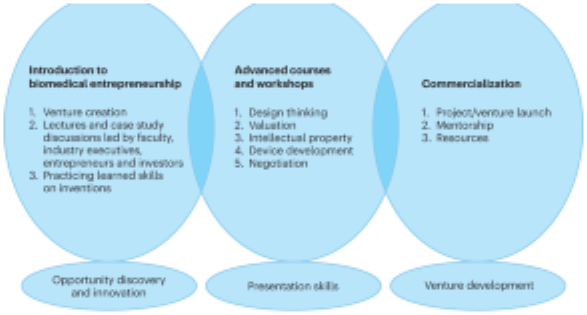


[*Nature Biotechnology*'s academic spinouts 2021]

Nature Biotechnology's annual survey highlights academic startups that are, among other things, designing circular RNA therapeutics, tackling cancer with arenaviruses, creating psychedelics without the trip, editing genes and cells in vivo, harnessing the power of autoantibodies and editing the epigenome.

- Michael Eisenstein
- Ken Garber
- Laura DeFrancesco

News Feature3 Nov 2022 [Nature Biotechnology]



[Longitudinal data support university-based biomedical entrepreneurship education programs]

Early-stage researchers can greatly benefit from education and skills development in venture creation and commercialization of discoveries.

- Gabriel Vizgan
- Zachary Hill-Whilton
- Gabrielle Gold-von Simson

Career Feature15 Mar 2023 Nature Biotechnology



Commercialisation and entrepreneurship – an academic viewpoint

We spoke to Professor Kylie Vincent – professor of inorganic chemistry at the University of Oxford, co-founder of HydRegen Ltd, and Academic Champion for Women in Entrepreneurship – about turning academic research into industrial products.

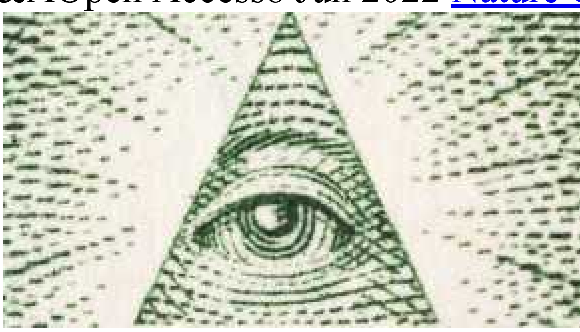
Q&AOpen Access21 Dec 2022 Nature Communications



A conversation on technology transfer

This Q&A about technology transfer is intended as a useful resource to the *Nature Communications* readership, particularly academic scientists working in the life and physical sciences who have an interest in commercializing their research. We spoke to Dr. Andrea Crottini, Head of the Technology Transfer Office at the École Polytechnique Fédérale de Lausanne, who provided insights into the possible avenues to pursue.

Q&AOpen Access8 Jun 2022 Nature Communications



Precision financing

Private biotech financing in 2023 is more discriminate, disciplined and demanding. Strong data and efficiency are de rigueur.

- Melanie Senior

News Feature12 Apr 2023 Nature Biotechnology



[The patenting versus publishing dilemma]

The process of patenting inventions may be complex. Academic researchers whose primary goal is getting their work published in scientific journals often face daunting doubts when it comes to understanding the interplay between publishing and patenting their findings. We asked Prof Frank Tietze questions from the perspective of academic researchers who wish to understand how the patenting process works and—most importantly—the relation between patenting and publishing.

Q&AOpen Access21 Mar 2023 Nature Communications



Is big tech draining AI talent from academia?

Movement to industry is raising concerns about the future researcher workforce and maintaining ethical expertise.

- Chris Woolston

Nature Index12 Oct 2022 Nature



[Why industry internships can be your ‘golden ticket’ to a prosperous career]

The three of us took a break from our PhD programmes for a stint that enriched our CVs and improved our chances of career success.

- Eric Juskewitz
- Kathleen Anne Heck
- Nancy Saana Banono

Career Column18 Mar 2021 Nature



How business-savvy scientists can find success in the risky start-up world

In the first of a three-part series on science start-ups, Nature Careers explores how scientists with a sound business idea can thrive as entrepreneurs, and why leaving academia isn’t required.

- Amber Dance

Career Feature13 Feb 2019 Nature



How the pandemic inspired a new generation of creators

The COVID-19 crisis echoed the ‘all-hands-on-deck’ response to the Second World War, but such agility needs to be maintained.

- Chris Woolston

Nature Index7 Dec 2022 Nature

This article was downloaded by **calibre** from https://www.nature.com/collections/cchcahhcab